HAL open science

# Stochastic differential equations : strong well-posedness of singular and degenerate equations; numerical analysis of decoupled forward backward systems of McKean-Vlasov type

Paul Éric Chaudru de Raynal



## ▶ To cite this version:

Paul Éric Chaudru de Raynal. Stochastic differential equations : strong well-posedness of singular and degenerate equations; numerical analysis of decoupled forward backward systems of McKean-Vlasov type. General Mathematics [math.GM]. Université Nice Sophia Antipolis, 2013. English. NNT : 2013NICE4115 . tel-00954418

## HAL Id: tel-00954418

## https://theses.hal.science/tel-00954418

Submitted on 2 Mar 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ NICE-SOPHIA ANTIPOLIS — UFR Sciences

École Doctorale Sciences Fondamentales et Appliquées

# THÈSE

pour obtenir le titre de

**Docteur en Sciences de l'Université Nice-Sophia Antipolis**

Discipline : Mathématiques

présentée et soutenue par

**Paul-Eric Chaudru de Raynal**

---

# Équations différentielles stochastiques : résolubilité forte d'équations singulières dégénérées ; analyse numérique de systèmes progressifs-rétrogrades de McKean-Vlasov

---

Thèse dirigée par **François Delarue**

soutenue le 6 décembre 2013

devant le jury composé de

| | | | |
|---|---|---|---|
| M. | Cédric Bernardin | Examinateur | Professeur des Universités |
| Mme. | Mireille Bossy | Examinateur | Directeur de Recherche INRIA |
| M. | Philippe Briand | Examinateur | Professeur des Universités |
| M. | François Delarue | Directeur | Professeur des Universités |
| M. | Antoine Lejay | Rapporteur | Directeur de Recherche INRIA |
| M. | Stéphane Menozzi | Rapporteur | Professeur des Universités |

Laboratoire Jean-Alexandre Dieudonné, Parc Valrose, 06108 Nice Cedex 2

# Merci

En démarrant cette thèse voilà un peu plus de trois ans, je ne me doutais pas à quel point cette expérience se révélerait enrichissante. Le corollaire de ce travail est une multitude de rencontres, tant mathématiques qu' humaines, et, bizarrement, une propension particulière au café, breuvage miraculeux aux multiples facettes qui me sauva plus d'une fois la mise. Voilà trois années tellement vite écoulées, où se sont mêlées joies, mathématiques, amitiés, doutes, soulagements et euphories. Trois années propices aux anecdotes en tout genre. Il y a forcément beaucoup de moments à rappeler, beaucoup de personnes à remercier. D'avance, et par peur d'en oublier, je remercie tous ceux que, de près ou de loin, j'ai pu croiser au cours de ces années et avec qui j'ai pu partager un petit quelque chose, ou un presque rien.

À l'origine de tout ceci il y a François Delarue. La chance m'a été offerte d'entreprendre mes premiers pas dans le monde de la recherche mathématique sous sa houlette, je lui en suis profondément reconnaissant. Je ne saurais, en si peu de lignes, lui témoigner toute ma gratitude. Pour ta disponibilité, ta patience, tes éclairages et suggestions, toutes ces heures passées ainsi que toutes les jolies mathématiques auxquelles tu m'as initié et que tu m'as faites découvrir : merci.

Merci à Antoine Lejay et Stéphane Menozzi d'avoir accepté de rapporter cette thèse. Je leur suis particulièrement reconnaissant, et suis sincèrement honoré qu'ils aient pris le temps de s'intéresser à mon travail. Les travaux de Stéphane Menozzi, en collaboration avec François Delarue, ont une place particulièrement importante dans mes recherches. Ma reconnaissance va aussi aux autres membres du jury : merci à Cédric Bernardin, Mireille Bossy et Philippe Briand d'avoir accepté de participer à ma soutenance.

J'ai particulièrement apprécié l'ambiance au laboratoire. Les membres, chercheurs chevronnés, en herbe et personnels techniques ont donné beaucoup de couleur à mes journées par leur gentillesse et leurs attentions. Mon nouveau voisin, Philippe Maisonobe, en est un exemple.

Ceux de l'équipe “Proba-Stats”, que je tiens à remercier. Plus particulièrement Yannick Baraud, Christine Malot, Sylvain Rubenthaleur, avec qui les discussions ont toujours été intéressantes. Merci à Patricia Reynaud-Bouret, deux fois, car j'ai aussi pu bénéficier de ton enthousiasme et ta générosité par connexité.

Ensuite, le “staff technique” du labo, tous à l'écoute et présents au besoin : Julien, Jean-Marc, Jean-Paul, Angélique, Francine et tous les autres. Merci de votre aide précieuse.

Puis les doctorants. Il est difficile d'imaginer vivier plus abondant d'étudiants sympathiques que le “LJAD”. Matheux, débatteurs, blablateurs ou un peu tout à la fois. Il y a ceux qu'on croise brièvement, qu'à peine arrivés ils sont déjà partis ; ceux qu'un café rapproche, avec lesquels on refait le monde : une fois deux fois, et encore et encore...
Les compagnons de “cellule”, premiers consultés quand on a oublié une démo où un résultat, quand on a besoin d'une traduction en anglais ou n'importe. Ceux qui sont passés, restés, arrivés, camarade d'exode du deuxième au troisième et de quelques nocturnes.

D'autres, ou les mêmes d'ailleurs, avec qui on a tapé des maths, plus ou moins, un surtout qui m'a fait manger du code des deux cotés de la Manche, même au pt'i déj ! J'espère que d'autres occasions se présenteront !

Les anciens, qui sont presque tous partis et qui ont donné une saveur toute particulière aux déjeuners de mes deux premières années. Extravagants, délirants, (choquants !). Ceux avec qui on a "tracé la route", et dont la voiture porte les stigmates. Ceux qui savent déguster un vrai café en débattant, pour le plaisir, de telle et telle question, parce que n'être pas d'accord c'est finalement s'entendre sur la nécessité de discuter. Ceux de Sophia et leurs D. facts. Les nouveaux, qui sont arrivés et qui ont pris le relais. Ceux du GDR GOT, attention, bossez bien, pour vous aussi "*winter is coming*". Le temps est passé tellement vite, dire que certains se sont même mariés !

J'ai une pensée pour tout ceux que j'ai rencontrés le long du chemin qui m'a conduit à la thèse : à Bourges, pour avoir partagé la même cigarette, le même vélo et le même centre-ville trop petit ; à Tours, prêts à batailler autour d'une roulée et d'un café à cinquante cents (non, quarante à l'IUT) ; à Nice, qui m'ont initié aux pratiques locales et qui s'en sont allés dans le nord ; à Paris, au "DEA" pour des "poz clopes" toujours intéressantes, et les débats météorologiques incessants.

Il y a ma famille, à qui je dois tellement, tout même. Savoir que je pouvais compter sur eux à tout moment m'a beaucoup aidé. Les chamailleries, fous rires, discussions ont été de vraies bouffées d'air pur. Et puis la famille s'est agrandie, et j'aimerai vous remercier pour votre accueil chaleureux et tout ce que vous m'avez fait, gaieté et bonne humeur sont toujours au rendez-vous avec vous !

Et il vaut mieux que je m'arrête ici, si jamais on veut commencer les maths.

*Après le bac, sans trop savoir pourquoi, j'ai suivi une petite lumière qui m'a emmené faire des maths. Quelques années plus tard j'en suis là, j'écris ces lignes et suis toujours ma belle Étoile. Merci d'avoir égayé toutes ces années et d'avoir accepté d'illuminer les prochaines. Sans toi, rien de tout cela n'aurait été possible.*

# Organisation du mémoire

La première partie de la thèse est introductive et rédigée en français. Elle permet de replacer les travaux effectués dans leur contexte et d'y décrire les principaux outils et arguments qui ont contribué à les mettre en forme. En conséquence, certaines notions basiques d'analyse stochastique sont exposées. Les résultats y sont ensuite énoncés et les preuves esquissées. Des perspectives sont finalement données.

Le reste du manuscrit se divise en deux parties regroupant les deux thèmes abordés durant cette thèse. Chacun de ces thèmes fait l'objet d'un chapitre rédigé en anglais :

(1) Le premier chapitre traite de la *résolubilité forte d'EDS dégénérées de dérive Hölder*. Il fait l'objet d'une pré-publication [CdR12] en révision aux *annales de l'IHP*.

(2) Le second chapitre est un travail effectué en collaboration avec Camilo Andrés García Trillos, et porte sur l'étude d'un *schéma numérique de type cubature pour les EDSR de McKean-Vlasov*. Il fait l'objet d'une pré-publication [CdRGT13] soumise à *Stochastic Processes and their applications*.

Finalement, puisque certaines des perspectives énoncées dans l'introduction, ou que certaines applications des résultats sont disponibles, on attache à chaque chapitre une annexe explicitant ces calculs. Ainsi :

(a) L'annexe A traite de la différentiabilité par rapport à la condition initiale du flot d'EDS dégénérées de dérive Hölder et bornée.

(b) L'annexe B porte sur la différentiabilité du semi-groupe non homogène et du flot de la solution d'équations différentielles stochastiques de McKean-Vlasov par rapport à la condition initiale.

Ces annexes sont elles aussi rédigées en anglais.

# Table des matières

# Introduction

## *0* Préliminaires : EDP et probabilités

*Les liens qui unissent l'analyse stochastique et celle des équations aux dérivées partielles sont au coeur des travaux présentés dans cette thèse. Cette section, purement introductive, vise à décrire brièvement ces liens et à en présenter certaines conséquences et applications. Le tout à évidemment été choisi de manière à introduire des outils utilisés pour la démonstration des résultats de ce mémoire, et la présentation n'est par conséquent pas exhaustive.*

### *0.1* Représentation des solutions d'une EDP

La *formule de Kolmogorov* constitue le lien fondamental entre EDP et probabilités. Dans le cadre de l'équation de la chaleur, la solution $u$ de :

$$\frac{\partial}{\partial t}u(t,x) - \frac{1}{2}\Delta u(t,x) = 0,\ \forall (t,x) \in \mathbb{R}^{+,*} \times \mathbb{R} \quad u(0,x) = \phi(x),$$

s'écrit :

$$u(t,x) = \mathbb{E}[\phi(x + B_t)],$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ muni d'une filtration naturelle $(\mathcal{F}_t)_{0\leq t\leq T}$.

Cette représentation peut aussi se comprendre de la manière suivante : en remontant le long de la trajectoire d'un mouvement brownien depuis un certain temps $t > 0$, la formule d'Itô nous dit que la solution $u$ est constante en espérance :

$$u(t-s, B_s) = u(t,x) + \int_0^s \frac{\partial}{\partial x}u(t-r, B_r)dB_r.$$

C'est une *martingale*. L'inversion de temps peut être contournée en considérant le problème *rétrograde* avec comme condition terminale $\phi$ à un certain temps $T > 0$, et dans ce cas la solution $u$ s'écrit

$$u(t,x) = \mathbb{E}[\phi(B_T^{t,x})],$$

où l'exposant $(t,x)$ indique les conditions initiales : $(B_s^{t,x},\ t \leq s \leq T)$ est un mouvement brownien partant à l'instant $t$ du point $x$. Plus généralement, si on note $\mathcal{L}$ un opérateur différentiel du second ordre :

$$\mathcal{L} := \sum_{i=1}^d b_i(t,x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j}, \tag{0.1}$$

où $b$ et $a$ sont des coefficients réguliers de $\mathbb{R}^+ \times \mathbb{R}^d$ dans $\mathbb{R}^d$ (resp. $\mathcal{M}_d(\mathbb{R})$) [1], et qu'on s'intéresse au *problème de Cauchy* [2] :

$$\frac{\partial}{\partial t}u(t,x) + \mathcal{L}u(t,x) = 0,\ \forall (t,x) \in [0,T) \times \mathbb{R}, \tag{0.2}$$

alors la solution régulière $u$ à ce problème s'écrit :

$$u(t,x) = \mathbb{E}[u(T, X_T^{t,x})].$$

Ici, la dynamique du processus $(X_s^{t,x},\ t \leq s \leq T)$ est donnée par l'*équation différentielle stochastique* (EDS) suivante :

$$dX_s^{t,x} = b(s, X_s^{t,x})ds + \sigma(s, X_s^{t,x})dB_s, \quad X_t^{t,x} = x, \tag{0.3}$$

où la matrice $\sigma$ est telle que $\sigma\sigma^* := a$. On remet au paragraphe suivant la discussion sur la résolubilité de cette équation.

Par ailleurs, dans le cas où l'EDP a un terme source $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ régulier, la solution s'écrit :

$$u(t,x) = \mathbb{E}\left[u(T, X_T^{t,x}) + \int_t^T f(s, X_s^{t,x})ds\right].$$

## *0.2* Solution d'une équation différentielle stochastique

***Résolubilité forte.*** La notion de solution forte d'une équation différentielle stochastique est très proche de celle d'une équation différentielle ordinaire. En particulier, sous des conditions de régularité Lipschitz des coefficients, on peut montrer l'existence d'une unique *solution forte*, c'est à dire un processus $(X_t,\ 0 \leq t \leq T)$ adapté à la filtration du brownien vérifiant (0.3).

A l'instar des équations différentielles ordinaires, il ne paraît pas évident de pouvoir s'affranchir des hypothèses de type Lipschitz sur les coefficients [3]. Par exemple, *l'équation de Tanaka* [4] n'admet pas de solutions fortes. *Peut-on néanmoins donner un sens à la résolubilité de cette équation ?*
La réponse est *oui*. En probabilité, il est usuel de ramener l'étude d'une variable aléatoire à celle de sa loi : c'est ce que l'on appelle la *résolubilité faible* d'une EDS.

***Résolubilité faible et problème de martingale.*** Une solution faible est un quintuplet $\big(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}, (B_t)_{t\geq 0}, (X_t^x)_{(t\geq 0)}\big)$. Dit autrement, le brownien et sa filtration font maintenant partie de la solution. La notion de solution faible est bien définie, puisque la résolubilité forte implique la résolubilité faible (voir [YW71]). La question est maintenant de savoir *quelles sont les conditions de résolubilité faible d'un système.*

La réponse est donnée par Stroock et Varadhan dans [SV79]. L'approche entreprise est ingénieuse : sur le modèle de la formule de Kolmogorov, le problème est réinterprété en un *problème de martingale*. Les auteurs montrent que les systèmes à dérives mesurables

1. L'ensemble des matrices réelles de taille $d \times d$.
2. ici formulé de manière rétrograde.
3. Tout du moins, pas sans contrepartie.
4. $dX_t^x = \text{signe}(X_t^x)dB_t, \quad X_0 = x$.

et bornées et à matrices de diffusions continues et uniformément non dégénérées[5] sont résolubles au sens faible. Pour la démonstration, un résultat d'analyse sur les *effets régularisants* des opérateurs elliptiques est nécessaire. Il s'agit d'un contrôle $\mathbb{L}_p$, $p > 1$ sur les dérivées d'ordre 2 de la solution de l'équation elliptique $\mathcal{L}u(x) = f(x), \ \forall x \in \mathbb{R}^d$, connu sous le nom d'inégalité de Calderòn-Zygmund.

*Le constat est frappant : c'est l'effet régularisant des opérateurs elliptiques qui permet d'assouplir les conditions de résolubilité faible d'une EDS.*

## *0.3* L'effet régularisant

Pour illustrer la notion d'effet régularisant, il faut revenir à la solution $u$ de l'équation de la chaleur : elle s'écrit comme la convoluée de la condition au bord $\phi$ avec la densité gaussienne. Pour peu que cette condition au bord soit mesurable et bornée, la solution est bornée et infiniment différentiable à dérivées bornées. C'est ce qu'on appelle *l'effet régularisant du noyau de la chaleur sur la donnée $\phi$.*

Cet effet subsiste dans le cas plus général d'un opérateur du type de $\mathcal{L}$ défini par (0.1) à coefficients réguliers et dont la matrice de diffusion est uniformément non-dégénérée. En fait, c'est l'hypothèse de non-dégénérescence qui est cruciale ici. Une belle illustration de l'effet régularisant d'opérateurs elliptiques dont les coefficients sont seulement bornés est le résultat de Krylov et Safonov obtenu en 1979 dans [KS79].

## *0.4* Probabilités de transitions.

Lorsque l'EDS est bien posée, la solution est *markovienne*. On peut alors définir la famille d'opérateurs $(P_{t,s})_{0\leq t\leq s}$ qui à toute fonction mesurable bornée $f$ de $\mathbb{R}^d$ dans $\mathbb{R}$ fait correspondre $P_{t,s}f(x) = \mathbb{E}[f(X_s^{t,x})]$, où $(X_s^{t,x}, \ t \leq s \leq T)$ est la solution de (0.3).

La propriété de Markov se comprend comme une propriété de semi-groupe qui, ainsi défini, est *Fellerien*. De plus, pour toute fonction $f$ dans $C_b^2(\mathbb{R}^d, \mathbb{R})$, on peut donner son générateur infinitésimal $\mathcal{G}$ :

$$\mathcal{G}_t f := \lim_{s\to 0} \frac{1}{s}(P_{t,t+s}f - f),$$

soit encore

$$\mathbb{E}[f(X_s^{t,x})] - f(x) = \int_t^s \mathbb{E}[\mathcal{G}_r f(X_r^{t,x})]dr, \tag{0.4}$$

et on identifie formellement $\mathcal{G}$ à l'opérateur $\mathcal{L}$ grâce à la formule d'Itô. Puisque la dynamique infinitésimale suffit à caractériser le processus, l'unique solution faible de l'EDS (0.3) est appelée *processus de diffusion de générateur $\mathcal{L}$.*

Par ailleurs, l'équation (0.4) permet de déduire l'évolution de la probabilité de transition du processus. Celle-ci est décrite par le problème parabolique progressif associé à $\mathcal{L}$ avec comme condition initiale une masse de Dirac :

$$\frac{d}{dt}\int \varphi(z)\mu_{0,t}(x,dz) = \int \mathcal{L}_t\varphi(z)\mu_{0,t}(x,dz), \quad \mu_{0,0}(x,\{x\}) = 1,$$

pour toute fonction $\varphi$ de $\mathbb{R}^d$ dans $\mathbb{R}$ régulière : c'est l'équation de *Fokker-Planck*.

5. *i.e.* $\exists\Lambda > 0$ tq $\forall(t,x) \in \mathbb{R}^+ \times \mathbb{R}^d, \ \forall y \in \mathbb{R}^d, \Lambda^{-1} \leq [a(t,x)y]\cdot[y] \leq \Lambda$, "$\cdot$" désignant le produit scalaire euclidien.

Lorsque cette probabilité admet une densité $p(t,x;s,y)dy := \mu_{t,s}(x,dy)$, celle-ci est la *solution fondamentale de l'opérateur* $\partial_t + \mathcal{L}$ : elle vérifie le problème parabolique associé à $\mathcal{L}$ avec comme condition au bord la masse de Dirac.

Dans ce cas, les solutions du problème de Cauchy s'obtiennent en intégrant la condition au bord contre la solution fondamentale de l'EDP.

En conséquence, la solution de l'EDP hérite de la régularité de la solution fondamentale. Lorsque celle-ci est régulière, l'opérateur est *hypoelliptique*.

### *0.5* Hypoellipticité

De manière concise, un opérateur du type $\mathcal{L}$ ou $\partial_t + \mathcal{L}$ est dit *hypoelliptique* si, pour toutes distributions $u$, $\mathcal{L}u$ est régulière implique que $u$ est régulière. Le lien étroit entre effet régularisant et propagation du bruit permet de comprendre qu'ellipticité et régularité des coefficients garantissent l'hypoellipticité.

Au contraire, obtenir un opérateur hypoelliptique alors que le bruit dégénère semble moins évident. Cependant, l'ellipticité n'est pas une condition nécessaire à l'hypoellipticité. En ce sens, *l'exemple de Kolmogorov* [Kol34] :

$$\mathcal{L}^K := \frac{1}{2}\Delta_{x_1} + \alpha x_1 \frac{\partial}{\partial x_2}, \tag{0.5}$$

constitue une très belle illustration d'un opérateur hypoelliptique dont la matrice de diffusion n'est pas elliptique. Lorsque $\alpha$ est non-nul, Kolmogrov a montré que $\mathcal{L}^K$ admettait une solution fondamentale $p^K$, gaussienne.

Cette constatation se généralise : c'est un résultat d'analyse très célèbre dû à Hörmander dans les années 60, [Hö67]. Soit $\{V_i,\ 0 \le i \le d\}$ une famille de champs de vecteurs réguliers sur $\mathbb{R}^d$. On introduit l'opérateur sous forme divergente $\mathcal{A}$,

$$\mathcal{A} := \sum_{i=1}^{d} V_0^i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{k,j=1}^{d} V_k^j \frac{\partial}{\partial x_j} \sum_{i=1}^{d} V_k^i \frac{\partial}{\partial x_i}, \tag{0.6}$$

a contrario des opérateurs non-divergents du type de (0.1). Lorsque les crochets de Lie [6] $\{V_i,\ i \ge 0, [V_{i_1}, V_{i_2}], i_1, i_2 \ge 0,\ [[V_{i_1}, V_{i_2}], V_{i_3}], , i_1, i_2, i_3 \ge 0, \cdots\}$, évalués en chaque point $x$ engendrent $\mathbb{R}^d$ tout entier, l'opérateur $\mathcal{A}$ est hypoelliptique.

Pour déduire de ce résultat les conditions d'existence d'une densité de transition du processus de diffusion de générateur $\mathcal{A}$, il faut appliquer le théorème sur $\mathbb{R}^+ \times \mathbb{R}^d$ à l'opérateur : $\partial_t + \mathcal{A}$. L'analogie avec (0.6) se fait en considérant la famille de champs de vecteurs $\{\bar{V}_0,\ V_i,\ 1 \le i \le d\}$ où $\bar{V}_0 := V_0 + \partial_t$.

### *0.6* Vers le(s) cas non linéaire(s)

Au delà de l'affaiblissement des conditions d'ellipticité, l'extension de la formule de Kolmogorov aux EDP non-linéaires a motivé de nombreux travaux. Ici, on distingue deux directions quant à la non-linéarité.

***Les EDP semi-linéaires et équations différentielles stochastiques rétrogrades.*** L'EDP dont la non-linéarité se concentre sur le terme source est dite semi-linéaire et s'écrit,

6. On rappelle que $[V_1, V_2] = V_1 \nabla_x V_2 - V_2 \nabla_x V_1$.

de manière générique :

$$\partial_t u + \mathcal{L}u(t,x) + f\left(t, x, u(t,x), \sigma\frac{\partial}{\partial x}u(t,x)\right) = 0,\ \forall (t,x) \in [0,T) \times \mathbb{R}, \quad u(T,x) = \phi(x).$$

Pour la représentation probabiliste, l'idée consiste à séparer les deux termes : on attribue la partie linéaire au processus de diffusion de générateur $\mathcal{L}$ et en appliquant de façon formelle la formule d'Itô à la solution $u$, on obtient pour le second processus la dynamique suivante :

$$Y_t := u(t, X_t) = \phi(X_T) + \int_t^T f(r, X_r, Y_r, Z_r)dr - \int_t^T Z_r dB_r, \tag{0.7}$$

où l'on a posé $(Z_t = \sigma(t, X_t)[\partial u/\partial x](t, X_t))_{t\geq 0}$. La fonction $f$, faisant intervenir la non-linéarité, est appelée le *générateur* de l'équation. C'est une EDS rétrograde (EDSR).

*A priori* le problème semble mal posé : il faut résoudre une équation à deux inconnues dont chaque membre semble appartenir à une tribu différente. C'est en fait le second problème qui permet de fermer l'équation : le processus $(Z_t)_{t\geq 0}$ permet de ramener le générateur et la condition terminale dans la bonne tribu. Une solution est un couple $(Y_t, Z_t)_{t\geq 0}$ *progressivement mesurable* et *vérifiant* (0.7).

On réfère aux travaux fondateurs de Pardoux et Peng [PP90, PP92] pour une exposition plus détaillée de ces méthodes.

***Les équations de McKean-Vlasov.*** On a vu que l'équation de Fokker-Planck décrivait la dynamique des probabilités de transition $(\mu_{0,t})_{t\geq 0}$ du processus de diffusion $(X_t)_{t\geq 0}$ de générateur $(\mathcal{L}_t)_{t\geq 0}$. Dans certains cas, cette famille de générateurs peut elle-même dépendre des probabilités de transition, l'évolution est alors donnée par l'équation :

$$\frac{\partial}{\partial t}\int \varphi(z)\mu_{0,t}(x,dz) = \int \mathcal{L}\big(\mu_{0,t}(x)\big)\varphi(z)\mu_{0,t}(x,dz), \quad \mu_{0,0}(x,\{x\}) = 1.$$

En conséquence, la dynamique du processus sous-jacent est elle aussi non-linéaire : les coefficients de dérive $b$ et de diffusion $\sigma$ dépendent des marginales du processus

$$dX_t = b(t, X_t, \mu_{0,t}(x))dt + \sigma(t, X_t, \mu_{0,t}(x))dB_t, \quad X_0 = x. \tag{0.8}$$

Lorsqu'il y a unicité, ce processus est markovien, où la propriété de Markov doit maintenant se comprendre sur l'espace augmenté $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ [7].

De tels processus apparaissent naturellement comme limite de système de particules en interaction. Les notes de Snitzman [Szn91] donnent un aperçu plus détaillé et bien plus vaste de la théorie.

## *0.7* Contributions de la thèse

La discussion entreprise dans les paragraphes précédents permet de situer le travail de cette thèse.

Le premier travail, présenté en section 1 et développé dans le Chapitre 1, traite de la *résolubilité forte* d'EDS dégénérées dont la dérive n'est pas lipschitzienne. La démonstration repose sur le lien étroit entre EDP et probabilités, à travers une étude par *parametrix* de l'*effet régularisant d'une EDP dégénérée*. A ce titre, les conditions *d'hypoellipticité*, qui

7. Ici $\mathcal{P}(\mathbb{R}^d)$ désigne l'espace des mesures de probabilité sur $\mathbb{R}^d$.

assurent l'existence de tels effets sont fondamentales. En particulier, la connexion avec *l'exemple de Kolmogorov* est au coeur de ce travail.

Le second, présenté en section 2 et développé dans le Chapitre 2, traite de l'étude de *schémas numériques* pour des équations différentielles stochastiques *progressives-rétrogrades de McKean-Vlasov*. Ces équations font intervenir les deux *non-linéarités* précédemment évoquées. L'algorithme obtenu peut être paramétré pour obtenir des vitesses d'approximation de la loi de la composante progressive de tout ordre et des vitesses d'ordre 1 et 2 pour la composante rétrograde. Cette procédure repose sur les *méthodes de cubature* sur l'espace de Wiener. Les *estimations de gradients* de la solution de l'EDP associée y jouent un rôle majeur.

## *1* Premier chapitre : résolubilité de systèmes différentiels aléatoires à dérive Hölder et bruit dégénéré

*Peut-on tirer parti de l'effet régularisant pour affaiblir les conditions de résolubilité forte ?* En 1974, dans [Zvo74], Zvonkin montre que les EDS uni-dimensionnelles à dérive seulement bornée et à matrice de diffusion uniformément non-dégénérée et lipschitzienne admettent une unique solution forte. Le résultat est ensuite étendu par Veretennikov [Ver80] au cas multi-dimensionnel, puis ce sont Krylov et Röckner [KR05] qui démontrent que les EDS à dérive bornée dans $\mathbb{L}_p$, $p > d$ et à matrice de diffusion unitaire sont résolubles au sens fort. Enfin, Zhang étend dans [Zha05] le résultat de Krylov et Röckner au cas d'une matrice de diffusion non-unitaire uniformément elliptique et Sobolev .

De manière synthétique, *la perturbation du système a permis de "restaurer la résolubilité"*[8] en dehors du cadre Cauchy-Lipschitz. A la lumière de la première section, on comprend que l'ajout de bruit permet de bénéficier de l'effet régularisant des opérateurs différentiels du second ordre à matrice de diffusion uniformément non-dégénérée.

Cette constatation a inspiré de nombreux auteurs. Par exemple, Fedrizzi et Flandoli donnent une nouvelle approche du résultat de Krylov et Röckner dans [FF11], et Flandoli, Gubinelli et Priola montrent dans [FGP10] que le constat s'étend au cas infini-dimensionnel en démontrant qu'une équation de transport perturbée par un bruit non-dégénéré est résoluble au sens fort. Tous les travaux dans cette direction ne sont pas cités, on réfère à [Fla11] et aux références qui s'y trouvent pour une description plus détaillée.

Dans ces travaux, l'hypothèse de non-dégénérescence du bruit est capitale. Ces résultats reposent en effet sur l'étude de l'EDP

$$\partial_t u(t,x) + \mathcal{L}u(t,x) = b(x),\ \forall (t,x) \in [0,T) \times \mathbb{R}^d,\ u(T,x) = 0.$$

Au vu des discussions antérieures sur l'hypoellipticité, il paraît raisonnable d'envisager l'extension du résultat à certains cadres dégénérés. La question posée dans ce chapitre est donc la suivante : *existe-il des EDS à dérive non lipschitzienne et dont le bruit dégénère qui sont résolubles au sens fort ?*

La réponse à la question est *oui*. Voici le résultat prouvé :

8. Évidemment, la notion de résolubilité a été modifiée.

**Théorème 1.1.** *Soit $T$ un réel positif et $d$ un entier naturel, on considère le système*

$$\begin{aligned} dX_t^1 &= F_1(t, X_t^1, X_t^2)dt + \sigma(t, X_t^1, X_t^2)dW_t, \\ dX_t^2 &= F_2(t, X_t^1, X_t^2)dt, \end{aligned} \tag{1.1}$$

*où $(W_t, t \geq 0)$ est un mouvement brownien $d$-dimensionnel défini sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ muni de la filtration naturelle $(\mathcal{F}_t)_{t\geq 0}$ ; $F_1$, $F_2$ sont des fonctions mesurables de $[0,T] \times \mathbb{R}^d \times \mathbb{R}^d$ dans $\mathbb{R}^d$ et $\sigma$ est une fonction mesurable de $[0,T] \times \mathbb{R}^d \times \mathbb{R}^d$ dans $\mathcal{M}_d(\mathbb{R})$ (l'ensemble des matrices réelles de taille $d \times d$).*

*On suppose que la matrice de diffusion $\sigma$ est uniformément elliptique et lipschitzienne en espace, uniformément en temps et que les coefficients de dérive $F_i$, $i = 1, 2$ sont de régularité Hölder en espace et d'exposant de Hölder strictement supérieur à $2/3$ pour la régularité par rapport à la deuxième variable. On suppose que le coefficient $F_2$ est dérivable par rapport à la première composante et que sa Jacobienne appartient à un sous ensemble convexe fermé des matrices inversibles. La dérivée est en outre supposée être de régularité Hölder.*

*Alors, il existe une unique solution forte au système (1.1).*

On peut, brièvement, résumer les hypothèses de ce résultat de la façon suivante. La condition de régularité Hölder des dérives est reliée à l'approche perturbative retenue pour démontrer ce résultat : le parametrix. L'hypothèse d'existence et d'inversibilité de la Jacobienne de $F_2$ garantit l'hypoellipticité du système, à l'instar de l'exemple de Kolmogorov présenté dans la section précédente. Enfin, le seuil critique de $2/3$ pour la régularité Hölder des dérives par rapport à la composante régulière intervient comme une "compensation" pour la dégénérescence du processus.

### *1.1* Heuristique

Afin de ne se concentrer que sur les mécanismes de preuve, et pour simplifier les calculs présentés, on se place dans le cas où $d = 1$, $F_1 = 0$, $\sigma = 1$ et $F_2(t, x_1, x_2) = f_2(x_2) + \alpha x_1$, $\alpha \in \mathbb{R}^*$. L'exposant de régularité Hölder de $f_2$ est noté $\beta_2$. Par ailleurs, on note $X = (X^1, X^2)^*$, $F = (F_1, F_2)^*$, $B = (1, 0)^*$ et $\mathcal{L}$ le générateur du processus de diffusion $X$.

L'heuristique est la suivante : on suppose que la fonction $u$, solution de l'EDP

$$\partial_t u(t,x) + \mathcal{L}u(t,x) = F(t,x), \text{ sur } [0,T) \times \mathbb{R}, \quad u(T,x) = 0_{\mathbb{R}^2},$$

et que son gradient $\nabla_x u$ sont lipschitziens en espace, avec une constante de Lipschitz $C_T$ telle que $C_T$ tende vers 0 avec $T$. Par ailleurs, un argument de compacité assure l'existence d'au moins une solution faible au système (1.1) (voir [SV79]). On peut, formellement, appliquer la formule d'Itô à $X_t - u(t, X_t)$ et on obtient que le processus $(X_t)_{0\leq t\leq T}$ se réécrit :

$$X_t^x = x - u(t,x) + u(t, X_t^x) - \int_0^t [B\nabla_x u(s, X_s^x) - B]\, dW_s. \tag{1.2}$$

Ainsi, si $(Y_t)_{0\leq t\leq T}$ est une autre solution de l'EDS, alors $(Y_t)_{0\leq t\leq T}$ admet une représentation du type de celle ci-dessus et on en déduit qu'il existe une constante positive $C_T$, tendant vers 0 avec $T$, telle que :

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}|X_t^x-Y_t^x|^2\right] &\le \mathbb{E}\left[\sup_{0\le t\le T}|u(t,X_t^x)-u(t,Y_t^x)|^2\right]\\ &\quad+\frac{1}{2}\int_0^T\mathbb{E}\left[|B\nabla_x u(s,X_s^x)-B\nabla_x u(s,Y_s^x)|^2\right]ds\\ &\le C_T\mathbb{E}\left[\sup_{0\le t\le T}|X_t^x-Y_t^x|^2\right]\\ &\quad+\frac{1}{2}C_T\mathbb{E}\left[\sup_{0\le t\le T}|X_t^x-Y_t^x|^2\right].\end{aligned}$$

S'en suit l'unicité forte pour $T$ "suffisamment petit". Il ne reste alors qu'à itérer le procédé...

Le résultat découle donc de la formule d'Itô et de l'analyse de l'EDP sous-jacente à la solution. Cette étude constitue le cœur de la preuve. Par ailleurs, on souligne qu'il n'est nécessaire d'obtenir une estimation lipschitzienne que sur $u$ et $B\nabla_x u = ([\partial u/\partial x_1], 0)^*$.

### *1.2* Étude de l'EDP.

On se replace dans les conditions de l'heuristique décrite dans la sous section précédente. Il s'agit d'étudier l'effet régularisant de l'opérateur

$$\mathcal{L} \;=\; \frac{1}{2}\frac{\partial^2}{\partial x_1^2}+(F_2(x_2)+\alpha x_1)\frac{\partial}{\partial x_2}. \tag{1.3}$$

Plus précisément, on cherche à exhiber une théorie forte pour la solution $u=(u_1,u_2)^*$ du système d'EDP :

$$\begin{aligned}&\partial_t u_i(t,x_1,x_2)+\mathcal{L}u_i(t,x_1,x_2)=F_i(t,x_1,x_2),\ \text{pour tout }(t,x)\in[0,T)\times\mathbb{R}^{2d},\\ &u_i(T,x_1,x_2)=0_{\mathbb{R}^d},\quad i=1,2,\end{aligned}\tag{1.4}$$

*i.e.* obtenir l'existence d'une solution $u$ telle que $u$ et $[\partial u/\partial x_1]$ satisfassent des estimations lipschitziennes en espace uniformément en temps. Le fait d'imposer à la solution de s'annuler sur le bord est important : cela permet l'obtention de constantes de Lipschitz petites en temps petit.

Une première remarque permet de simplifier le cadre de travail. Il n'est en réalité pas nécessaire de démontrer l'existence d'une solution classique à (1.4) : par une procédure de régularisation, il suffit d'exhiber les estimations dans un cadre régulier, mais ne dépendant que des paramètres apparaissant dans les hypothèses du Théorème 1.1. On montre en fait que

**Proposition 1.1.** *Il existe une constante positive $C_T$ ne dépendant que de paramètres énoncés dans le Théorème 1.1 et de $T$, telle que*

$$\left\|\frac{\partial}{\partial x_1}u^n\right\|_\infty+\left\|\frac{\partial}{\partial x_2}u^n\right\|_\infty+\left\|\frac{\partial^2}{\partial x_1^2}u^n\right\|_\infty+\left\|\frac{\partial^2}{\partial x_1\partial x_2}u^n\right\|_\infty\le C_T \tag{1.5}$$

*où l'exposant "n" fait référence à la procédure de régularisation et* où la constante $C_T$ peut être choisie aussi petite que voulu, pourvu que l'intervalle considéré soit petit. *Cette propriété est essentielle pour la démonstration.*

En combinant ces estimations à l'heuristique et en passant à la limite sur la procédure de régularisation, on obtient le résultat voulu. La démonstration de la Proposition 1.1, qui constitue le véritable défi, se fait grâce à un *développement en série parametrix au premier ordre de la solution.*

La méthode parametrix est une méthode perturbative. Initiée par Levi dans les années 1900, l'application qu'on en fait peut aussi se voir comme un développement de McKean et Singer [MS67] au premier ordre.

L'idée de base de la méthode provient essentiellement de la constatation suivante : "*en temps petit, la densité de transition de la solution d'une EDS à coefficients variables réguliers est proche de celle à coefficients gelés*[9]". Dans le cas classique, où la matrice de diffusion est uniformément non-dégénérée, la solution à coefficients constants est un *processus gaussien.* Dit autrement, il s'agit de considérer le générateur comme *une perturbation du Laplacien.* Le choix du point de gel des coefficients est très important. La perturbation doit être de l'ordre des trajectoires typiques du processus de diffusion ayant pour générateur l'opérateur gelé.

L'étude se divise ainsi en deux parties :

1./ Une étude du système gelé, essentielle à la mise en œuvre du parametrix.

2./ Une partie technique où chacune des dérivées de la solution est calculée et estimée.

Pour clarifier l'exposition, on oublie volontairement l'exposant "$n$" faisant référence à la procédure de régularisation dans la suite de la section.

*1./ Le système gelé.* Dans ce contexte dégénéré, il faut tout d'abord s'assurer de l'existence d'une densité au processus gelé. Lorsque $f_2 = 0$, on se retrouve dans le cas de l'exemple de Kolmogorov, et une condition pour obtenir l'existence d'une densité est alors $\alpha \neq 0$ : le bruit injecté dans la première composante doit pouvoir être propagé dans la seconde (ici *via* la dérive). Il paraît ainsi naturel d'étendre l'hypothèse $\alpha \neq 0$ au système (1.1). C'est le sens de l'hypothèse d'inversibilité de la Jacobienne de $F_2$.

La condition additionnelle imposant à la Jacobienne de $F_2$ d'appartenir à un sous ensemble convexe fermé des matrices inversibles est technique et reliée à l'extension du résultat de Kolmogorov, due à Delarue et Menozzi dans [DM10]. On renvoie à la section 3 et à l'exemple 3.5 de leur article pour plus de précisions.

Le choix du point de gel des coefficients est lui aussi délicat. Dans le cas elliptique, la décroissance hors de la diagonale est en l'exponentielle du carré de la distance entre le point d'arrivé et de départ. Le gel des coefficients se fait donc en un de ces points. Ici, la dégénérescence du bruit modifie les trajectoires typiques du processus. Pour le comprendre il faut à nouveau revenir à l'exemple de Kolmogorov et en particulier à la forme de sa densité $p^K$ donnée par :

$$p^K(0, x_1, x_2; t, y_1, y_2) = \frac{\sqrt{3}}{\alpha \pi t^2} \exp\left(-\frac{1}{2}\left|K_t^{-1/2}(y_1 - x_1, y_2 - x_2 - \alpha t x_1)^*\right|^2\right), \tag{1.6}$$

où la matrice de covariance est :

$$K_t := \begin{pmatrix} t & (1/2)\alpha t^2 \\ (1/2)\alpha t^2 & (1/3)\alpha^2 t^3 \end{pmatrix}. \tag{1.7}$$

9. Dont les coefficients sont pris en un point donné.

C'est l'échelle du système qu'il faut regarder : chacune des deux composantes du système vit à une échelle propre, de l'ordre de $t^{i-1/2}$ (penser, par exemple, à un mouvement brownien et son intégrale itérée en temps). Il en résulte que le transport de la condition initiale de la première composante dans la seconde, de l'ordre de $t$, n'est pas négligeable en temps petit.

Ainsi, et c'est ce que suggère la forme de l'exponentielle de la solution fondamentale (1.6) de l'exemple de Kolmogorov, il faut geler la première variable au point de départ $x_1$ et la seconde le long du transport $(\theta_{t,s})_{t\leq s\leq T}$ de la condition initiale de la première composante dans la seconde. Dans le cas de l'heuristique, ce transport est assuré par la solution l'équation [10] :

$$\frac{d}{ds}\theta_{t,s}(\xi_2) = f_2(\theta_{t,s}(\xi_2)) + \alpha x_1, \quad \theta_{t,t}(\xi_2) = \xi_2,\ \xi_2 \in \mathbb{R}^d,$$

avec comme condition initiale $\xi_2 = x_2$.

On en déduit alors le système gelé

$$\begin{cases} d\tilde{X}^1_s = dW_s, & \tilde{X}^1_t = x_1, \\ d\tilde{X}^2_s = \left(f_2(\theta^2_{t,s}(\xi_2)) + \alpha\tilde{X}^1_s\right)ds, & \tilde{X}^2_t = x_2, \end{cases} \tag{1.8}$$

dont la densité $\tilde{p}$ et ses dérivées admettent des bornes gaussiennes :

**Lemme 1.2.** *Pour $T$ suffisamment petit, il existe deux constantes positives $C$ et $c$ telles que, quels que soient $(t, x_1, x_2)$ et $(s, y_1, y_2)$ dans $[0, T] \times \mathbb{R} \times \mathbb{R}$*

$$\tilde{p}(t, x_1, x_2; s, y_1, y_2) \leq C\hat{p}_c(t, x_1, x_2; s, y_1, y_2),$$

*où*

$$\begin{aligned}&\hat{p}_c(t, x_1, x_2; s, y_1, y_2)\\ &= \frac{c}{(s-t)^{2d}}\exp\left(-c\left(\frac{|y_1 - x_1|^2}{s-t} + \frac{|y_2 - x_2 - \int_t^s f_2(\theta_{t,r}(\xi_2)dr - (s-t)\alpha x_1)|^2}{(s-t)^3}\right)\right),\end{aligned}$$

*et pour tout $s$ dans $[t, T]$ et tout entier $N_1, N_2$ inférieur où égal à deux*

$$\begin{aligned}&\left|\frac{\partial^{N_1}}{\partial x_1^{N_1}}\frac{\partial^{N_2}}{\partial x_2^{N_2}}\tilde{p}(t, x_1, x_2; s, y_1, y_2)\right|\\ &\qquad \leq C(s-t)^{-[3N_2+N_1]/2}\hat{p}_c(t, x_1, x_2; s, y_1, y_2).\end{aligned} \tag{1.9}$$

On souligne que ces bornes ont une décroissance diagonale en puissance $1/2$ dans la direction diffusive et $3/2$ dans la direction dégénérée.

*2./ Représentation en parametrix et différentiation.* Dans un deuxième temps, il s'agit de donner une représentation adéquate de la solution. En effectuant un développement en série parametrix au premier ordre, l'équation (1.4) est réécrite

$$\begin{aligned}&\partial_t u_i(t, x_1, x_2) + \tilde{\mathcal{L}}u_i(t, x_1, x_2) = -(\mathcal{L} - \tilde{\mathcal{L}})u_i(t, x_1, x_2) + F_i(t, x_1, x_2),\\ &u_i(T, x_1, x_2) = 0_{\mathbb{R}^d}, \quad i = 1, 2,\end{aligned}$$

où $\tilde{\mathcal{L}}$ est le générateur du processus gelé (1.8). On obtient alors une représentation de la solution $u$ comme convolution temps-espace (notée $\otimes$) du noyau $(\mathcal{L} - \tilde{\mathcal{L}})u$ et du terme source $F$ avec la densité du processus gelé :

$$u(t, x_1, x_2) = -(\mathcal{L} - \tilde{\mathcal{L}})u \otimes \tilde{p}(t, x_1, x_2; ., ., .) + F \otimes \tilde{p}(t, x_1, x_2; ., ., .), \tag{1.10}$$

10. bien posée dans le cadre régularisé.

soit,

$$
\begin{aligned}
u_i(t,x_1,x_2) \quad = \quad & \int_t^T \int_{\mathbb{R}^{2d}} F_i(s,y_1,y_2)\tilde{p}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds \\
& - \int_t^T \int_{\mathbb{R}^{2d}} (f_2(y_2)-f_2(\theta_{t,s}(\xi_2)))\frac{\partial u_i}{\partial x_2}(s,y_1,y_2)\tilde{p}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds,
\end{aligned} \tag{1.11}
$$

pour $i=1,2$. On ne présente que l'analyse de $u_2$, celle de $u_1$ pouvant être conduite de la même manière. Par ailleurs, on rappelle que $F_2(t,x_1,x_2)=f_2(x_2)+\alpha x_1$.

Pour obtenir les bornes souhaitées, il faut dériver cette représentation et estimer les normes infinies des dérivées. Afin de simplifier la lecture, on suppose qu'il est possible d'intervertir intégrale et dérivation.

On remarque alors par l'application du Lemme 1.2 que chaque dérivation de la solution génère une singularité en temps qui n'est pas forcément intégrable. En particulier, la dérivée première dans la direction dégénérée, les dérivées secondes dans la direction diffusive et croisée génèrent des singularités en temps d'ordre respectivement $3/2$, 1 et 2. L'idée, afin de régulariser cette singularité, est de tirer parti de la *décroissance gaussienne* de la densité de transition.

On applique cet argument dans le "pire" des cas, à savoir celui de la dérivation croisée. Grâce à un *argument de centrage*, on peut écrire :

$$
\begin{aligned}
& \frac{\partial^2}{\partial x_1 \partial x_2} u_2(t,x_1,x_2) \\
& = \int_t^T \int_{\mathbb{R}^{2d}} (f_2(y_2)-f_2(\theta_{t,s}(\xi_2)))\frac{\partial^2}{\partial x_1 \partial x_2}\tilde{p}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds \\
& \quad - \int_t^T \int_{\mathbb{R}^{2d}} (f_2(y_2)-f_2(\theta_{t,s}(\xi_2)))\frac{\partial u_2}{\partial x_2}(s,y_1,y_2)\frac{\partial^2}{\partial x_1 \partial x_2}\tilde{p}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds.
\end{aligned} \tag{1.12}
$$

Ainsi, en utilisant le Lemme 1.2 on obtient que

$$
\begin{aligned}
& \left|\frac{\partial^2}{\partial x_1 \partial x_2} u_2(t,x_1,x_2)\right| \\
& \leq C\left(1+\left\|\frac{\partial}{\partial x_2}u_2\right\|_\infty\right)\int_t^T\int_{\mathbb{R}^{2d}}\left|\frac{y_2-\theta_{t,s}(\xi_2)}{(s-t)^{3/2}}\right|^{\beta_2}(s-t)^{-2+3\beta_2/2}\hat{p}_c(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds,
\end{aligned} \tag{1.13}
$$

et en posant $\xi_2=x_2$ et en intégrant en espace, on a

$$
\left|\frac{\partial^2}{\partial x_1 \partial x_2} u_2(t,x_1,x_2)\right| \leq C\left(1+\left\|\frac{\partial}{\partial x_2}u_2\right\|_\infty\right)\int_t^T (s-t)^{-2+3\beta_2/2}ds, \tag{1.14}
$$

et le membre de droite n'est intégrable que si $\beta_2 > 2/3$ : l'exposant critique donné en énoncé.

On remarque que la représentation implicite de la solution fait apparaître la norme infinie de la dérivée $[\partial u_2/\partial x_2]$ dans la borne (1.14). Pour contourner ce problème, il faut ensuite estimer cette quantité puis injecter l'estimation obtenue dans la borne (1.14). Cet argument circulaire se généralise au cas du système énoncé dans le Théorème 1.1.

### *1.3* Généralisation

On se place dans le cas de l'heuristique mais où $F_1(t, y_1, y_2) = f_1(y_1)$ avec $f_1$ une fonction $\beta_1$-Hölder et bornée. Si on gèle ce coefficient en la valeur de départ [11] "$x_1$" alors le processus gelé admet encore une densité de transition et les estimations du Lemme 1.2 restent vraies. On obtient comme représentation pour $u$ :

$$\begin{aligned} u_i(t, x_1, x_2) &= -\int_t^T \int_{\mathbb{R}^{2d}} (f_1(y_1) - f_1(\xi_1)) \frac{\partial u_i}{\partial x_1}(t, y_1, y_2) \tilde{p}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds \\ &\quad - \int_t^T \int_{\mathbb{R}^{2d}} (f_2(y_2) - f_2(\theta_{t,s}(\xi_2))) \frac{\partial u_i}{\partial x_2}(s, y_1, y_2) \tilde{p}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds \\ &\quad + \int_t^T \int_{\mathbb{R}^{2d}} F_i(s, y_1, y_2) \tilde{p}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds, \end{aligned}$$

pour $i = 1, 2$ et où $\xi_1 \in \mathbb{R}^d$. On se concentre à nouveau sur l'analyse de $u_2$, celle de $u_1$ se déduisant encore des même arguments.

La différence avec (1.12) est le premier terme du membre de droite faisant intervenir une dépendance non-linéaire en la première variable. Cette dépendance va modifier la gestion de la singularité engendrée par les dérivations $[\partial/\partial x_2]$ et $[\partial^2/\partial x_2 \partial x_1]$ et la rendre plus délicate.

En effet, en recourant à la même analyse que celle effectuée précédemment et en posant $\xi_1 = x_1$, on obtient une intégrale supplémentaire du type de (1.11) avec comme intégrande :

$$s \in (t, T] \mapsto \left| \frac{y_1 - \xi_1}{(s-t)^{1/2}} \right|^{\beta_1} (s-t)^{-2+\beta_1/2} \hat{p}_c(t, x_1, x_2; s, y_1, y_2). \tag{1.15}$$

Or, cette fonction n'est intégrable sur l'intervalle $(t, T]$ que pour $\beta_1 > 2$. Ce problème est dû à la décroissance hors de la diagonale de la première composante qui n'est que de l'ordre de $1/2$. Pour compenser la singularité d'ordre 2, il faut donc recourir à d'autres arguments.

L'astuce consiste à tirer parti de la présence de la solution dans la représentation en centrant l'intégrande autour de la dérivée de la solution :

$$\begin{aligned} &\frac{\partial^2}{\partial x_1 \partial x_2} \int_t^T \int_{\mathbb{R}^{2d}} (f_1(y_1) - f_1(\xi_1)) \frac{\partial u_2}{\partial x_1}(t, y_1, y_2) \tilde{p}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds \\ &= \int_t^T \int_{\mathbb{R}^{2d}} \left\{ \Big( f_1(y_1) - f_1(\xi_1) \Big) \left( \frac{\partial u_2}{\partial x_1}(t, y_1, y_2) - \frac{\partial u_2}{\partial x_1}(t, y_1, \theta_{t,s}(\xi_2)) \right) \right. \\ &\qquad \left. \times \frac{\partial^2}{\partial x_1 \partial x_2} \tilde{p}(t, x_1, x_2; s, y_1, y_2) \right\} dy_1 dy_2 ds, \end{aligned}$$

et puisque :

$$\left| \frac{\partial}{\partial x_1} u_2(t, y_1, y_2) - \frac{\partial}{\partial x_1} u_2(t, y_1, \theta_{t,s}(\xi_2)) \right| \leq \left\| \frac{\partial^2}{\partial x_1 \partial x_2} u_2 \right\|_\infty |y_2 - \theta_{t,s}(\xi_2)|,$$

11. L'échelle de la première composante étant diffusive.

on en déduit, en utilisant la décroissance gaussienne de la densité et en posant $(\xi_1, \xi_2) = (x_1, x_2)$, que :

$$\left| \frac{\partial^2}{\partial x_1 \partial x_2} \int_t^T \int_{\mathbb{R}^{2d}} (f_1(y_1) - f_1(x_1)) \frac{\partial u_2}{\partial x_1}(t, y_1, y_2) \tilde{p}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds \right|$$
$$\leq C ||f_1||_{\beta_1} \left\| \frac{\partial^2}{\partial x_1 \partial x_2} u_2 \right\|_\infty \int_t^T (s-t)^{(-1+\beta_1)/2} ds,$$

où $||f_1||_{\beta_1}$ est la semi-norme de Hölder de $f_1$. En combinant cette estimation avec les arguments précédents on obtient que

$$\left| \frac{\partial^2}{\partial x_1 \partial x_2} u_2(t, x_1, x_2) \right| \leq C_T \left( 1 + \left\| \frac{\partial}{\partial x_2} u_2 \right\|_\infty + \left\| \frac{\partial^2}{\partial x_1 \partial x_2} u_2 \right\|_\infty \right).$$

La constante $C_T$ pouvant être choisie aussi petite que voulue (pourvu que $T$ soit suffisamment petit) un argument circulaire permet de d'obtenir une estimation ne dépendant que de la norme infinie de $\partial u / \partial_{x_2}$. Pour généraliser cet argument cas du système du Théorème 1.1, où la dépendance en la première composante apparaît dans la dérive $F_2$, il faut estimer la régularité de $x_2 \mapsto [\partial u / \partial x_2](\cdot, \cdot, x_2)$. Cette estimation (en semi-norme de Hölder) est particulièrement délicate puisque chaque dérivation génère une singularité d'ordre $3/2$ (voir la section 1.4.4 du chapitre 1 pour plus de détails).

### *1.4* Perspectives

(1) Extension au cas où la dérive de la composante diffusive est bornée dans $\mathbb{L}_p$. En mettant en perspective ce travail avec ceux de Veretenikov [Ver80] et Krylov et Röckner [KR05], il semble naturel de supposer le coefficient de dérive de la composante diffusive borné dans $\mathbb{L}_p$, où $p$ est strictement supérieur à la dimension $d$ du processus.

En effet, dans l'article [Ver83], Veretenikov étudie le cas d'une équation du type (1.1) où la dérive de la composante diffusive est seulement bornée. En supposant suffisamment de régularité sur les coefficients par rapport à la composante dégénérée, l'auteur montre qu'il y a encore unicité forte pour le système.

L'idée est de traiter séparément les deux variables. En conséquence, on étudie deux EDP : une de type équation de la chaleur, qui offre de bonnes propriétés de régularisation, et une dégénérée, pour laquelle les coefficients sont supposés être réguliers. Bien qu'il ne s'intéresse pas à la régularisation du bruit transmis par la dérive, ce travail donne la marche à suivre pour analyser séparément les deux composantes, contrairement à l'approche faite dans ce manuscrit.

En outre, Krylov et Röckner montrent que dans le cas uniformément elliptique le système admet une unique solution forte pour un coefficient de dérive uniquement $\mathbb{L}_p$, $p > d$, ce qui laisse présager que cette hypothèse devrait être suffisante. La condition $p > d$ est une conséquence de la formule d'Itô-Krylov [Kry69] puisque, dans ce cas, l'EDP associée n'admet pas de solution au sens classique (seulement de type Sobolev).

(2) Seuil critique pour la régularité Hölder de la dérive. Bien que le seuil critique de $2/3$ pour l'exposant de Hölder soit cohérent avec l'approche exposée, le comportement intrinsèque du système suggère un exposant de $1/3$.

Comme il a été exposé, le seuil de $2/3$ provient de l'estimation en norme infinie des dérivées secondes de la solution de l'EDP. Il n'est cependant pas indispensable d'obtenir de petites constantes ou d'avoir des estimations en norme infinie pour ces dérivées afin de mettre en oeuvre l'heuristique exposée au début de la section.

En fait, le terme faisant intervenir les dérivées secondes pourrait être majoré en utilisant une inégalité de Krylov [Kry09] couplée à l'estimation ad hoc de Calderòn et Zygmund [12]. Néanmoins, il n'a pas été possible de trouver ce type d'estimation dans la littérature : bien que Bramanti et Zu aient récemment proposé des estimations de type Calderòn et Zygmund pour des opérateurs hypoelliptiques [BZ11], l'extension de l'estimation de Krylov à des processus d'Itô dégénérés, même hypoelliptiques, demeure une question ouverte.

A titre d'application, réussir à abaisser le seuil de l'exposant critique à $1/3$ permettrait d'envisager l'étude de la résolubilité forte pour une chaîne d'oscillateurs toute entière, du type de celle étudiée dans [DM10].

(3) Différentiabilité du flot de la solution par rapport à la condition initiale. Dans [FGP10], les auteurs montrent que le système non-dégénéré à coefficient de dérive singulier définit un flot de difféomorphisme. La même question s'étend à ce cadre dégénéré.

On montre en annexe A du premier chapitre que le système (1.1) définit lui aussi un flot différentiable en la condition initiale. La stratégie consiste à achever l'étude de l'EDP (1.4) en montrant qu'elle admet une unique solution classique. Ensuite, il s'agit d'appliquer la démarche usuelle de Kunita [Kun82] au processus changé de dérive (1.2) pour montrer que le flot ainsi défini est différentiable.

(4) Estimation des dérivées de la densité de transition du processus. Dans [DM10], Delarue et Menozzi obtiennent des estimations gaussiennes supérieures et inférieures de la densité du processus (1.1) pour des dérives lipschitzienne et une diffusion hölderienne.

Leur méthode ne permet cependant pas d'obtenir des estimations de la dérivée de la densité. Le fait d'obtenir des estimations sur les gradients de la solution de l'EDP dégénérée permet d'envisager l'obtention d'estimations gaussiennes des dérivées de la densité du processus (1.1) dan le cas lipschitzien.

Essentiellement, ils s'agirait de combiner l'approche de Delarue et Menozzi et celle précédemment exposée pour obtenir une borne sur la décroissance diagonale de la densité. L'estimation hors de la diagonale se déduirait ensuite d'une méthode d'interpolation sur le modèle de la preuve issue de l'ouvrage de Stroock [Str08], Section 3.

## *2* Deuxième chapitre : un algorithme de type cubature pour la simulation de processus progressifs et progressifs-rétrogrades découplés de McKean-Vlasov

L'extension de la formule de Kolmogorov aux cas non-linéaires a permis l'exploration de nombreux axes de recherche. Plus récemment, les travaux sur la théorie des jeux à champ moyen de Lasry et Lions [LL06b, LL06b, LL07] ont amené certains auteurs à s'intéresser au couplage des deux non-linéarités (la semi-linéarité et la non-linéarité de McKean-Vlasov). Le second travail de cette thèse s'inscrit dans cette direction : il s'agit de donner un algorithme probabiliste pour la simulation d'équations différentielles stochastique progressives-rétrogrades découplées de McKean-Vlasov (EDSPR de McKean-Vlasov).

L'algorithme présenté est basé sur les méthodes de cubature, récemment introduites par Lyons et Victoir [LV04], et se décompose en deux étapes :

– Une première étape de construction d'un arbre déterministe approchant la loi de la composante progressive. Cet arbre peut être paramétré de manière à obtenir n'importe quel ordre d'approximation (en terme de pas de discrétisation de l'intervalle).

12. Un contrôle $\mathbb{L}_p$ sur les dérivées secondes en fonction du terme source.

– Une seconde étape permettant d'approcher la composante rétrograde avec des ordres d'approximation de 1 et 2.

De manière générique, les EDSPR découplées de McKean-Vlasov s'écrivent :

$$\begin{cases} dX_t^x = \sum_{i=0}^d V_i(t, X_t^x, \mu_t) \circ dB_t^i \\ dY_t^x = -f(t, X_t^x, Y_t^x, Z_t^x, \mu_t^{X,Y})dt + Z_t^x dB_t^{1:d} \\ X_0^x = x, \quad Y_T^x = \phi(X_T^x), \end{cases} \tag{2.1}$$

définis sur un intervalle $[0,T]$, $T$ un réel strictement positif donné où $(W_t^{1:d}, t \geq 0)$ est un mouvement brownien $d$-dimensionnel défini sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ muni de la filtration naturelle $(\mathcal{F}_t)_{t\geq 0}$ et où $W_t^0 = t$, les champs de vecteurs $V_i$ sur $\mathbb{R}^d$ sont supposés être réguliers [13] tout comme le générateur $f$, et $\phi$ est une fonction mesurable de $\mathbb{R}^d$ dans $\mathbb{R}$. Ici, $\mu$ (resp. $\mu^{X,Y}$) désigne la loi du processus progressif (resp. couple processus progressif - rétrograde). On adopte une écriture Stratonovitch des équations, signalée par "$\circ$".

## *2.1* Contexte

C'est la théorie des jeux à champ moyen, introduite par Lasry et Lions dans [LL06b, LL06b, LL07] et dans une série de cours au collège de France (voir les notes qui en sont issues [Car10]), et notamment le pendant probabiliste de ces jeux, développé par Carmona et Delarue [CD12b, CD12a] qui a motivé ce travail.

En effet, le système (2.1) s'inscrit dans une classe plus large d'équations, appelées équations progressives rétrogrades de type champ moyen, qui s'écrivent de façon générique :

$$\begin{cases} dX_t^x = \sum_{i=0}^d V_i(t, X_t^x, Y_t, \mu^{X,Y}) \circ dB_t^i \\ dY_t^x = -f(t, X_t^x, Y_t^x, Z_t^x, \mu_t^{X,Y})dt + Z_t^x dB_t^{1:d} \\ X_0^x = x, \quad Y_T^x = \phi(X_T^x). \end{cases} \tag{2.2}$$

Ces équations sont les dérivées probabilistes des EDP étudiées par Lasry et Lions. On se contentera de ne décrire que brièvement les idées mises en oeuvre et on réfère aux travaux sus-cités pour une exposition plus détaillée.

On appelle jeu un système de $M$ joueurs interagissant. Chaque joueur à une structure de type diffusion donnée et identique [14] (les joueurs peuvent permuter), qu'ils peuvent contrôler par le biais de leur *stratégie*. On associe à chacun des joueurs une fonctionnelle d'utilité, de même forme, dépendant de la stratégie. Le but du jeu consiste à optimiser cette utilité.

Sous de bonnes hypothèses (rationalité des joueurs,...), la théorie des jeux dit qu'il existe un *équilibre de Nash* [15] : un état du jeu (un panel de contrôles) qui rend tout autre état sous optimale, au sens où, si on est dans cet équilibre et si un des joueurs tend à s'écarter de sa stratégie, alors il sous optimisera son utilité. La stratégie optimisant l'utilité du $M$ème joueur s'obtient ainsi au vu des ($M$-1)ème autres stratégies, et la fonctionnelle obtenue est disponible comme solution d'une équation d'un système d'EDP, de type Hamilton-Jacobi-Bellman (HJB).

On comprend que, pour un grand nombre de joueurs, le problème devient rapidement incalculable. L'idée proposée par Lasry et Lions consiste à approcher les équilibres du

13. Le sens donné à la régularité des coefficients par rapport à la composante Mckean-Vlasov sera précisé au fur et à mesure.

14. La forme est la même, mais les contrôles dont ils dépendent sont différents.

15. Le problème de l'existence d'un équilibre de Nash peut être contourné, en changeant d'ensemble des stratégies, voir la sous section 2.3 de [Car10].

système d'un grand nombre de joueurs par l'équilibre du système limite [16]. La terminologie de champ moyen prend alors tout son sens : les joueurs ne ressentent pas les dynamiques des autres joueurs mais seulement leurs distributions statistiques, à l'instar des systèmes de particules en interaction champ moyen en physique.

Les équations (des EDP) résultant de ces jeux sont nouvelles, au sens où sont couplées : une équation de type HJB, décrivant la dynamique de la fonctionnelle d'utilité optimale, et une équation de Kolmogorov, décrivant la dynamique de la distribution des joueurs. La nouveauté provient, entre autre, du fait que ces deux équations sont en "sens inverse" (l'équation de Kolmogorov est progressive, celle d'HJB rétrograde). Carmona et Delarue montrent que la dérivation probabiliste du système HJB peut encore se faire *via* une adaptation du principe du maximum stochastique. Le système qui en résulte est donné par (2.2).

**Le système** (2.1)**.** Le but du second chapitre est de proposer une première et nouvelle méthode probabiliste pour l'approximation des solutions d'EDSPR de type McKean-Vlasov intervenant dans la modélisation d'un problème de contrôle dans un environnement de type champ moyen.

En effet, le cadre abordé dans ce manuscrit est plus restreint que celui des jeux à champ moyen : le système (2.2) est couplé quand le système (2.1) fait intervenir des dynamiques découplées, sur le modèle de celles présentées dans la sous section sur les EDSPR.

Pour comprendre le problème de contrôle abordé ici il suffit de se replacer dans le contexte des jeux à champ moyen décris dans le paragraphe précédent. On suppose cette fois-ci que seul un joueur, marqué, est contrôlé et qu'il cherche à maximiser une certaine utilité. L'environnement est composé du système des autres joueurs, qui sont supposés ne pas ressentir l'action du joueur marqué.

Lorsque ce nombre de joueur est grand, on approche le système par le système limite et l'environnement devient de type champ moyen. Le découplage de la partie McKean-Vlasov des équations est assuré par l'inaction du joueur marqué sur les autre joueurs. A nouveau, la fonctionnelle d'utilité est disponible comme la solution d'une EDP de type HJB.

### *2.2* Préliminaires : méthodes de Cubature sur l'espace de Wiener.

Les méthodes de cubature de Lyons et Victoir [LV04] sont à la base de l'algorithme du chapitre 2. On se propose d'en donner une description détaillée car elle est nécessaire à la compréhension de l'algorithme et des idées mises en œuvre. On parle de $m \in \mathbb{N}^*$ cubature sur l'espace de Wiener dans le cas suivant :

**Définition**. *Étant donné un entier naturel $m$ et un réel positif $t$, une $m$-cubature sur l'espace de Wiener (l'ensemble des fonctions continues de $[0,t]$ à valeurs dans $\mathbb{R}^d$) est une mesure de probabilité discrète $\mathbb{Q}_t$ à support fini dans l'ensemble des fonctions continues à variations bornées de $[0,t]$ dans $\mathbb{R}^d$ telle que les espérances des intégrales du mouvement brownien de Stratonovitch itérées d'ordre $m$ soient les mêmes sous la mesure de cubature $\mathbb{Q}_t$ et sous la mesure de Wiener $\mathbb{P}$.*

Ainsi, à la place d'une EDS (sous forme Stratonovitch) de type :

$$dX_t = \sum_{i=0}^{d} V_i(X_t) \circ dB_t^i,$$

16. Et à démontrer que cette approche donne effectivement un optimum approché.

on résout un système d'équations différentielles pondérées, où le mouvement brownien est "remplacé" par une trajectoire continue à variations bornées $\omega_j : [0,t] \to \mathbb{R}^d$ à laquelle est associée un poids $\lambda_j$, $j \in \{1,\dots,n\}$, où $n = \mathrm{card}\{\mathrm{supp}(\mathbb{Q}_t)\}$ :

$$d\hat{X}_t^j = V_0(\hat{X}_t^j)dt + \sum_{i=1}^{d} V_i(\hat{X}_t^j)dw_j^i(t), \quad j = 1,\dots,n,$$

où on a adopté la convention $w_j^0(t) = t$. On comprend naturellement que ces méthodes permettent d'approcher l'espérance des fonctions régulières des solutions d'EDS par des polynômes. En utilisant la définition d'une mesure de cubature et un développement de Taylor, on en déduit que pour toute fonction régulière $F$ :

$$(\mathbb{E}_{\mathbb{P}} - \mathbb{E}_{\mathbb{Q}_t})[F(X_t)] \le Ct^{(m+1)/2} \max_{j \le m+2} \left\| \frac{\partial^j}{\partial x^j} F \right\|_\infty, \tag{2.3}$$

en majorant l'espérance du reste du développement de Taylor stochastique, voir par exemple l'ouvrage [KP92].

Bien sûr, si la constante $C$ ou le temps $t$ ne sont pas assez petits, l'erreur commise peut être grande. Afin de profiter au maximum de la méthode, il faut tirer parti de la structure markovienne du processus : au lieu d'appliquer une cubature sur l'intervalle $[0,T]$ tout entier, on l'applique à chaque pas d'une subdivision $T_0 = 0 < T_1 < \cdots < T_N = T$ de cet intervalle.

Cette approche aboutit à la construction d'un arbre dont chaque nœud possède un nombre $n$ (dépendant du degré de cubature $m$) de descendants. Si on note $\mathcal{S}_n(k)$ l'ensemble des multi-indices à valeur dans $\{1,\cdots,n\}$ de taille exactement $k$, alors chaque nœud de l'arbre à l'étape $k$ est indicé par un multi-indice $\pi \in \mathcal{S}_n(k)$.

A chaque nœud $\pi$ correspond la valeur d'une trajectoire $\omega_\pi$ prise au $k$ème instant de la subdivision. En outre, à chacun de ces nœuds est attaché le poids $\Lambda_\pi = \prod_{j_l \in \pi} \lambda_{j_l}$.

La construction des mesures de cubature sur chaque intervalle se fait déduit des propriétés d'échelles du mouvement brownien : étant donnée une mesure de cubature $\mathbb{Q}_1$ d'ordre $m$ de support $\{\omega_1,\dots,\omega_n\}$ et poids $\{\lambda_1,\dots,\lambda_n\}$ on peut en déduire une mesure de cubature $\mathbb{Q}_{t,t+h}$ sur l'intervalle $[t,t+h]$ d'ordre $m$ à support fini dans l'ensemble des fonctions continues à variations bornées de $[t,t+h]$ dans $\mathbb{R}^d$ donnée par $\{\tilde{\omega}_1,\cdots,\tilde{\omega}_n\}$ où $\tilde{\omega}_j : s \in [t,t+h] \mapsto \tilde{\omega}(s) = \sqrt{h}\omega_j((s-t)/h)$ pour tout $1 \le j \le n$ et dont la famille de poids reste $\{\lambda_1,\dots,\lambda_n\}$.

Ces méthodes exigent beaucoup de régularité, dans le sens où la fonctionnelle $F$ considérée doit pouvoir être développée en un polynôme de degré $m+2$. La pertinence de la méthode dépend donc de son extension au cas de fonctionnelles moins régulières. À nouveau, c'est l'effet régularisant de l'EDP sous-jacente qui joue un rôle clé dans cette extension.

A chaque instant $T_k$ de la subdivision on sait, par la propriété de Markov, que la cubature s'applique à la fonction $u(T_k,x) = \mathbb{E}[F(X_T)|X_{T_k} = x]$, solution du problème parabolique associé au générateur $\mathcal{L}$ de $X$ avec comme condition terminale $F$. Lorsque la matrice $V$ [17] est elliptique, que les coefficients sont réguliers et que $F$ est lipschitzienne, cette fonction est infiniment différentiable à dérivées bornées hors du bord : elle satisfait

$$\forall n \in \mathbb{N}^*,\ \exists C > 0 \text{ telle que } \frac{\partial^n}{\partial x^n} u(t,x) \le C||F||_{Lip}(T-t)^{(1-n)/2}, \tag{2.4}$$

17. La matrice réelle de taille $d \times d$ dont la $i$ème colonne est le vecteur $V_i$, $i = 1,\dots,d$.

où $||F||_{Lip}$ est la constante de Lipschitz de $F$, voir par exemple [Fri08] [18]. Ainsi, il suffit d'adapter la taille de la subdivision afin de bénéficier de cet effet : celle-ci peut être choisie suffisamment grossière loin du bord, puisque la solution est très régulière et plus fine près du bord, lorsque la solution devient moins régulière [19].

### *2.3* Algorithmes

La recherche d'un schéma numérique pour les solutions du système (2.1) a abouti à trois algorithmes :

(i) Un algorithme d'approximation de la loi de la composante progressive, sous forme d'un arbre de particules dont les dynamiques sont déterministes.

(ii) Deux algorithmes d'approximation de la solution de l'équation rétrograde d'ordre respectivement 1 et 2 (en terme de discrétisation temporelle). Ces deux algorithmes sont conditionnés à la donnée de l'arbre (i).

La clé de voûte, aussi bien des algorithmes que de l'analyse des erreurs, réside dans la constatation suivante qui sera utilisée de manière intensive :

*Etant donnée la loi du processus* (2.1)*, le système se comporte comme un système classique paramétré à coefficients inhomogènes.* ($\mathcal{C}$)

***2.3.1 Arbre de cubature : le cas de la composante progressive.*** Étant donnée une subdivision $T_0 = 0 < T_1 < \ldots < T_N = T$ de l'intervalle, une cubature d'ordre $m$ de support $\{\omega_1, \ldots, \omega_n\}$ [20] et de poids $\{\lambda_1, \ldots, \lambda_n\}$, il s'agit de construire une famille de mesures de probabilités discrètes $(\hat{\mu}_{T_k})_{0\leq k\leq N}$ approchant les marginales du processus $(X_t)_{0\leq t\leq T}$ aux instants $T_0, T_1, \ldots, T_N$. Cette mesure est donnée par les sommes pondérées (par des poids $\Lambda_\cdot$) de Dirac prisent en les points $\hat{X}^{\cdot}$ d'un arbre, noté $\mathcal{T}$.

Pour la construction de l'arbre $\mathcal{T}$, on décide de tirer parti au maximum de la constatation ($\mathcal{C}$). On procède donc comme suit : au début de chaque temps de discrétisation, on se donne une loi de probabilité (sur l'espace considéré). Cette loi étant gelée, le processus "redevient" la solution usuelle d'une EDS, et on peut appliquer l'algorithme de cubature sur l'intervalle discrétisé, il ne reste qu'à itérer le procédé... Évidemment, le choix de la loi est crucial pour assurer la convergence de l'algorithme. Un candidat naturel pour l'intervalle $[T_k, T_{k+1})$ est la mesure discrète engendrée par les solutions des EDO associées à la cubature obtenue à l'étape précédente : $\hat{\mu}_{T_k}$.

L'analogie avec la discussion sur les méthodes de cubature permet de comprendre que la taille de l'arbre dépend du cardinal du support de la mesure de cubature. On notera donc l'arbre $\mathcal{T}(m)$, où $m$ rappelle l'ordre de la cubature. L'algorithme est le suivant :

18. La condition d'ellipticité peut être réduite à une condition UFG [KS84, KS85, KS87], qui dégrade néanmoins ces bornes, voir [Nee11] pour une présentation.

19. Au sens où les bornes de ses gradients successifs sont de plus en plus grandes.

20. On rappelle que $\omega$ est la notation générique pour une trajectoire continue à variations bornées et on réfère au paragraphe sur les méthodes de cubature pour les notations.

**Algorithm 1** Arbre $\mathcal{T}(m)$

1: On initialise $(X_0, \Lambda_0, \hat{\mu}_{T_0}) = (x, 1, \delta_x)$.
2: A chaque étape $k \in \{1, \ldots, N-1\}$

a : On définit pour chaque noeud $\pi$ dans $\mathcal{S}_n(k)$ (l'ensemble des multi-indices de longueur $k$ et à valeur dans $\{1, \ldots, n\}$) et pour tout $j$ dans $\{1, \ldots, n\}$ le point $\hat{X}^{(\pi,j)}_{T_{k+1}}$ comme la valeur terminale de la solution de l'EDO

$$d\hat{X}^{(\pi,j)}_t = \sum_{i=1}^d V_i(t, \hat{X}^{(\pi,j)}_t, \hat{\mu}_{T_k}) \sqrt{\Delta_{T_{k+1}}} d\omega^i_j((t-T_k)/\Delta_{T_{k+1}}),$$
$$\hat{X}^{(\pi,j)}_{T_k} = \hat{X}^{\pi}_{T_k}$$

et on lui associe un poids $\Lambda_{(\pi,j)} = \Lambda_\pi \lambda_j$.

b : On pose $\hat{\mu}_{T_{k+1}} = \sum_{\pi \in \mathcal{S}_n(k+1)} \Lambda_\pi \delta_{\hat{X}^{\pi}_{T_{k+1}}}$.

Dans le chapitre 2, on quantifie l'action de la mesure $\mu_T$, issue de l'arbre $\mathcal{T}(m)$, sur n'importe quelle fonction régulière :

**Théorème 2.1.** *Si les champs de vecteur $V_i$ sont de régularité Lipschitz en la composante McKean-Vlasov pour la distance de Wasserstein-1*[21]*, alors il existe une constante positive $C$ telle que pour toute fonction $\phi$ régulière :*

$$\langle \hat{\mu}_T - \mu_T, \phi \rangle \leq C N^{-1/2} \sup_{j \leq m+1} ||\nabla^j_x \phi||_\infty, \tag{2.5}$$

*où $\langle \mu, \phi \rangle$ est la notation duale de $\int \phi d\mu$ et où $\mu_T$ est définie par l'algorithme 1. Si les champs de vecteur $V_i$ sont de régularité Lipschitz en la composante McKean-Vlasov pour la distance $d_{\mathcal{F}}$ entre deux mesure de probabilité $\nu$ et $\nu'$ sur $\mathbb{R}^d$ donnée par*

$$d_{\mathcal{F}}(\nu, \nu') = \sup_{\varphi \in \mathcal{F}} |\langle \varphi, \nu' - \nu \rangle|, \tag{2.6}$$

*où $\mathcal{F} = \{$des fonctions de $\mathbb{R}^d$ dans $\mathbb{R}$, bornées et infiniment différentiables à dérivées bornées$\}$, alors il existe une constante positive $C$ telle que pour toute fonction $\phi$ régulière :*

$$\langle \hat{\mu}_T - \mu_T, \phi \rangle \leq C N^{-1} \sup_{j \leq m+1} ||\nabla^j_x \phi||_\infty, \tag{2.7}$$

*où la mesure $\mu_T$ est définie par l'algorithme 1 ci dessus.*

Il faut souligner que les ordres d'approximation du Théorème 2.1 sont restreint à 1/2 ou 1, alors que l'approximation par cubature d'une EDS peut-être choisie aussi fine que voulue, pourvu que le paramètre "$m$" de la cubature soit choisi suffisamment grand. En regardant attentivement l'algorithme 1, on peut deviner que cette restriction est due au gel temporel de la loi injectée dans les coefficients de la solution de l'EDO à l'étape 2-a, et donc à une erreur de "type Euler"[22].

Dans ce cas, une approche classique pour augmenter l'ordre du schéma consiste à développer les coefficients. Ainsi, dans le cas particulier d'une dépendance du type

$$V_i(x, \mu_t) = V_i(x, \langle \varphi_i, \mu_t \rangle), \tag{2.8}$$

21. Pour rappel, la distance de Wasserstein 1 entre deux mesures de probabilité $\nu$ et $\nu'$ sur $\mathbb{R}^d$ est donnée par : $d(\nu, \nu') = \sup\{\langle \nu - \nu', \psi \rangle,\ \psi : \mathbb{R}^d \to \mathbb{R}$ 1-Lipschitz$\}$

22. Forte ou faible, suivant les conditions de régularité des coefficients

où $w \in \mathbb{R} \mapsto V_i(.,w)$ est bornée infiniment différentiable à dérivées bornées et où $\varphi_i$ est une fonction régulière de $\mathbb{R}^d$ dans $\mathbb{R}$, pour $i = 1, \ldots, d$, on peut effectuer un développement du coefficient à n'importe quel ordre $q-1$. On obtient alors un arbre paramétré par l'ordre de ce développement $\mathcal{T}(q,m)$ :

**Algorithm 2** Arbre $\mathcal{T}(q,m)$

1: On initialise pour tout $i$ dans $\{1,\ldots,d\}$ $F_i(t,\hat{\mu}_{T_0}) = \sum_{p=0}^{q-1} \frac{1}{p!}(t-T_0)^p \langle \delta_x, (\mathcal{L}^{\delta_x})^p \varphi_i \rangle$

2: A chaque étape $k \in \{1;\ldots,N-1\}$

  a : Idem que pour l'algorithme 1 en injectant le terme $F_i$ dans le coefficient $V_i$.

  b : Idem que pour l'algorithme 1.

  c : Pour tout $i$ dans $\{1,\ldots,d\}$, on calcule $F_i(t,\hat{\mu}_{T_{k+1}}) = \sum_{p=0}^{q-1} \frac{1}{p!}(t-T_k)^p \langle \hat{\mu}_{T_{k+1}}, (\mathcal{L}^{\hat{\mu}})^p \varphi_i \rangle$

où $\mathcal{L}^{\delta_x}$ (resp. $\mathcal{L}^{\hat{\mu}}$)) désigne le générateur dont la composante McKean-Vlasov des coefficients est gelée en $\delta_x$ (resp. $\hat{\mu}_{T_k}$ à l'étape $k$). On montre dans le chapitre 2 que

**Théorème 2.2.** *Si les champs de vecteur $V_i$ satisfont à (2.8), alors il existe une constante positive $C$ telle que pour toute fonction $\phi$ régulière :*

$$\langle \hat{\mu}_{T_k} - \mu_{T_k}, \phi \rangle \leq C N^{(m-1)/2 \wedge q} \sup_{j \leq m+1} ||\nabla_x^j \phi||_\infty, \tag{2.9}$$

*où la famille $(\mu_{T_k})_{0 \leq k \leq N}$ est définie par l'algorithme 2 ci dessus.*

Il faut noter que, dans ce cadre, c'est un contrôle sur l'approximation des marginales du processus pour *tous* les instants $T_k$, $k \in \{1,\ldots,N\}$ de la subdivision que l'on obtient. Dans la suite, on prend comme convention que l'arbre $\mathcal{T}(m)$ s'écrit $\mathcal{T}(1,m)$.

***2.3.2 Algorithme pour l'équation rétrograde.*** Une fois construit l'arbre $\mathcal{T}$, l'algorithme de simulation de la composante rétrograde peut être mis en œuvre. La forme spécifique de ces processus, brièvement abordée dans le paragraphe sur les EDSPR, donne la marche à suivre : la condition terminale étant connue, il s'agit de remonter l'arbre en calculant les valeurs approchées $(\hat{Y}, \hat{Z})$ de la solution du système rétrograde $(Y,Z)$ à chaque nœud comme une espérance conditionnelle.

Trois ingrédients sont donc nécessaires à la mise en œuvre de cette procédure : il faut tout d'abord se donner un schéma de discrétisation de la composante rétrograde, ensuite, il s'agit d'approcher la partie McKean-Vlasov du processus et, enfin, il faut calculer les espérances conditionnelles. A nouveau, la stratégie consiste à tirer profit du gel $(\mathcal{C})$ de la loi : *étant donnée une mesure, la solution de l'équation rétrograde s'écrit comme une fonction de la composante progressive*, comme dans le paragraphe sur les EDSPR, il existe deux fonctions $u$ et $v$ telles que :

$$Y_t = u(t,X_t),\ Z_t = V(t,X_t)\frac{\partial u}{\partial x}(t,X_t) =: v(t,X_t),$$

où $V$ est la matrice réelle de taille $d \times d$ ayant pour $i$ème colonne le champ $V_i$, $i = 1,\ldots,d$.

Ainsi, les espérances conditionnelles sont naturellement données par les mesures de cubature locales (de chaque nœud de l'arbre) et les composantes McKean-Vlasov se déduisent directement de l'arbre. On obtient alors un ordre d'approximation des espérances conditionnelles régi par l'ordre de la cubature classique et un ordre d'approximation de la composante McKean-Vlasov régi par l'ordre des arbres $(\mathcal{T}(q,m))_{q\geq 1,\ m\geq 1}$ définis par l'algorithme 1 ou 2. Reste la donnée d'un schéma.

Deux schémas sont proposés, basés sur un schéma de discrétisation de Zhang [Zha04], et tous deux conditionnellement à la donnée d'un arbre $\mathcal{T}$ :

- Le premier, d'ordre 1, peut-être comparé à une version explicite du schéma introduit par Crisan et Manolarakis dans [CM12] :

**Algorithm 3** Algorithme d'ordre 1 pour l'équation rétrograde

1: On initialise pour $\pi \in \mathcal{S}_n(N)$ : $\hat{Y}^\pi_{T_N} =: \hat{u}^1(T_N, X^{(\pi)}_{T_N}) = \phi(\hat{X}^\pi_{T_N})$ et $\hat{Z}^\pi_{T_N} =: \hat{v}^1(T_N, X^{(\pi)}_{T_N}) = 0$

2: A chaque étape $k$ à rebours dans $\{0,\ldots,N-1\}$, pour chaque $\pi$ dans $\mathcal{S}_n(k)$ :

a : On pose $\hat{v}^1(T_k, \hat{X}^\pi_{T_k}) = \dfrac{1}{\Delta_{T_{k+1}}} \displaystyle\sum_{j=1}^n \lambda_j \hat{u}^1(T_{k+1}, X^{(\pi,j)}_{T_{k+1}}) \sqrt{\Delta_{T_{k+1}}}\omega_j(1)$

b : Pour tout $j$ dans $\{1,\ldots,d\}$ on calcule $\hat{\Theta}^{\pi,1}_{k+1,k}(j) = \left(T_{k+1}, \hat{X}^{(\pi,j)}_{T_{k+1}}, \hat{u}^1(T_{k+1}, \hat{X}^{(\pi,j)}_{T_{k+1}}), \hat{v}^1(T_k, \hat{X}^\pi_{T_k}), \hat{\mu}^{X,Y}_{T_{k+1}})\right)$

c : On pose $\hat{u}^1(T_k, \hat{X}^\pi_{T_k}) = \displaystyle\sum_{j=1}^n \lambda_j \left(\hat{u}^1(T_{k+1}, X^{(\pi,j)}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\hat{\Theta}_{k+1,k}(j))\right)$

On montre dans le chapitre 2 que l'algorithme 3 permet de recouvrir les même ordres que pour la composante progressive :

**Théorème 2.3.** *Soit $\mathcal{T}(1,m)$ où $m \geq 3$ un arbre défini par l'algorithme 1. Si les champs de vecteur $V_i$ et le générateur $f$ sont lipschitziens en la composante McKean-Vlasov pour la distance de Wasserstein-1 alors pour toute fonction $\phi$ régulière, il existe une constante positive $C$ telle que :*

$$\max_{k\leq N,\ \pi\in\mathcal{S}_n(k)} |u(T_k, \hat{X}^\pi_{T_k}) - \hat{u}^1(T_k, \hat{X}^\pi_{T_k})| + \Delta^{1/2}_{T_k} |v(T_k, \hat{X}^\pi_{T_k}) - \hat{v}^1(T_k, \hat{X}^\pi_{T_k})| \leq CN^{-1/2}, \quad (2.10)$$

*où les quantités $\hat{u}^1$ et $\hat{v}^1$ sont définies par l'algorithme 3. De plus, si les champs de vecteur $V_i$ et le générateur $f$ sont lipschitziens en la composante McKean-Vlasov pour la distance $d_{\mathcal{F}}$ définie*[23] *par (2.6), alors pour toute fonction $\phi$ régulière, il existe une constante positive $C$ telle que :*

$$\max_{k\leq N,\ \pi\in\mathcal{S}_n(k)} |u(T_k, \hat{X}^\pi_{T_k}) - \hat{u}(T_k, \hat{X}^\pi_{T_k})| + (T_{k+1} - T_k)^{1/2} |v(T_k, \hat{X}^\pi_{T_k}) - \hat{v}(T_k, \hat{X}^\pi_{T_k})| \leq CN^{-1}, \quad (2.11)$$

*où les quantités $\hat{u}^1$ et $\hat{v}^1$ sont définies par l'algorithme 3.*

Il faut souligner que ce premier schéma est d'ordre 1. Ainsi, il n'est pas utile de préciser la version correspondant à une dépendance spécifique du type (2.8) puisque

23. Il faut adapter la classe $\mathcal{F}$ de fonctions dans le cas du générateur, la dépendance étant en la loi jointe, voir le corollaire 2.2.3 dans le chapitre 2.

ce cas est inclus dans la deuxième assertion du Théorème 2.3.

- Le second schéma est d'ordre 2. Il est basé sur un argument de *prédiction / correction* : un premier proxy du processus est construit, sur le modèle du premier schéma, afin d'en construire un second de manière "implicite". Cette procédure permet de gagner un ordre de convergence supplémentaire sans devoir calculer de dérivées.

  Pour ce schéma, on se place dans le cas où la famille $\{V_i,\ 1 \leq i \leq d\}$ satisfait la dépendance (2.8) et on suppose que le générateur $f$ vérifie le même type d'hypothèse :
$$f(\cdot,\cdot,\cdot,\cdot,\mu_t^{X,Y}) = f(\cdot,\cdot,\cdot,\cdot,\langle \mu_t^{X,Y}, \varphi_f(\cdot,\cdot)\rangle) \tag{2.12}$$
  où $w \in \mathbb{R} \mapsto f(\cdot,\cdot,\cdot,\cdot,w)$ est bornée infiniment différentiable à dérivées bornées et où $\varphi_f$ de $\mathbb{R}^d \times \mathbb{R}$ dans $\mathbb{R}$ est régulière.

  Le schéma est le suivant :

**Algorithm 4** Algorithme d'ordre 2 pour l'équation rétrograde

1: On initialise pour $\pi \in \mathcal{S}_n(N)$ : $\hat{Y}^\pi_{T_N} =: \hat{u}^2(T_N, X^{(\pi)}_{T_N}) = \phi(\hat{X}^\pi_{T_N})$ et $\hat{Z}^\pi_{T_N} =: \hat{v}^2(T_N, X^{(\pi)}_{T_N}) = 0$

2: On pose, pour $\pi \in \mathcal{S}_n(N-1)$ : $\hat{u}^2(T_{N-1}, \hat{X}^\pi_{T_{N-1}}) = \hat{u}^1(T_{N-1}, \hat{X}^\pi_{T_{N-1}})$, $\hat{v}^2(T_{N-1}, \hat{X}^\pi_{T_{N-1}}) = \hat{v}^1(T_{N-1}, \hat{X}^\pi_{T_{N-1}})$ et $F^2(T_{N-1}, \hat{\mu}_{T_{N-1}}) = \langle \hat{\mu}_{T_{N-1}}, \varphi_f(\cdot, \hat{u}^2(T_{N-1},\cdot))\rangle$

3: A chaque étape $k$ à rebours dans $\{0,\dots,N-2\}$ :

a : Pour tout $j$ dans $\{1,\dots,d\}$ on calcule :
$$\hat{\Theta}^{\pi,2}_{k+1}(j) = \left(T_{k+1}, \hat{X}^{(\pi,j)}_{T_{k+1}}, \hat{u}^2(T_{k+1}, \hat{X}^{(\pi,j)}_{T_{k+1}}), \hat{v}^2(T_{k+1}, \hat{X}^{(\pi,j)}_{T_{k+1}}), F^2(T_{k+1}, \hat{\mu}_{T_{k+1}})\right)$$
$$\hat{\zeta}^{(\pi,j)}_{k+1} := 4\frac{1}{\Delta_{T_{k+1}}}\sqrt{\Delta_{T_{k+1}}}\omega_j(1) - 6\frac{1}{\Delta^2_{T_{k+1}}}\int_{T_k}^{T_{k+1}} (s-T_k)\sqrt{\Delta_{T_{k+1}}}d\omega_j((s-T_k)/\Delta_{T_{k+1}})$$

b : On pose $\hat{v}^2(T_k, \hat{X}^\pi_{T_k}) = \sum_{j=1}^{\kappa} \lambda_j \left(\hat{u}^2(T_{k+1}, X^{(\pi,j)}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\hat{\Theta}^{(\pi,j),2}_{k+1})\right)\hat{\zeta}^{(\pi,j)}_{k+1}$

*Prediction* $\tilde{u}(T_k, \hat{X}^\pi_{T_k}) = \sum_{j=1}^{\kappa} \lambda_j \left(\hat{u}^2(T_{k+1}, X^{(\pi,j)}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\hat{\Theta}^{(\pi,j),2}_{k+1})\right)$

c : On pose $\tilde{F}(T_k, \hat{\mu}_{T_k}) = \langle \hat{\mu}_{T_k}, \varphi_f(\cdot, \tilde{u}(T_k,\cdot))\rangle$

*Correction* $\tilde{\Theta}^\pi_k = \left(T_k, \hat{X}^\pi_{T_k}, \tilde{u}(T_k, \hat{X}^\pi_{T_k}), \hat{v}^2(T_k, \hat{X}^\pi_{T_k}), \tilde{F}(T_k, \hat{\mu}_{T_k})\right)$

d : $\hat{u}^2(T_k, \hat{X}^\pi_{T_k}) = \sum_{j=1}^{\kappa} \lambda_j \left(\hat{u}^2(T_{k+1}, X^{(\pi,j)}_{T_{k+1}}) + \frac{1}{2}\Delta_{T_{k+1}}\left(f(\hat{\Theta}^{(\pi,j),2}_{k+1}) + f(\tilde{\Theta}^\pi_k)\right)\right)$

e : On pose $F^2(T_k, \hat{\mu}_{T_k}) = \langle \hat{\mu}_{T_k}, \varphi_f(\cdot, \hat{u}^2(T_k,\cdot))\rangle$

On montre que

**Théorème 2.4.** *Soit $\mathcal{T}(q,m)$ où $m \geq 5$ et $q \geq 2$ un arbre défini par l'algorithme 2. Si les champs de vecteur $V_i$ et le générateur $f$ satisfont (2.8) et (2.12), alors pour toute fonction $\phi$ régulière il existe une constante positive $C$ telle que :*
$$\max_{k\leq N,\ \pi\in\mathcal{S}_n(k)} |u(T_k, \hat{X}^\pi_{T_k}) - \hat{u}^2(T_k, \hat{X}^\pi_{T_k})| + (T_{k+1}-T_k)^{1/2}|v(T_k, \hat{X}^\pi_{T_k}) - \hat{v}^2(T_k, \hat{X}^\pi_{T_k})| \leq CN^{-2}, \tag{2.13}$$
*où les quantités $\hat{u}^2$ et $\hat{v}^2$ sont données par l'algorithme 4.*

Ces résultats sont illustrés dans le cas de dépendance du type (2.12) en section 2.3 du chapitre 2.

**Remarque.** *On peut objecter que les algorithmes présentés, puisqu'ils reposent sur la construction d'un arbre, ont une complexité exponentielle et qu'en comparaison les méthodes basées sur des systèmes de particules en interaction (voir [Bos05] pour une description et des références dans le cas de la composante progressive) sont moins coûteuses.*

*Cependant, il faut noter que les approximations obtenues ici permettent de bénéficier de la régularité des coefficients et sont aussi fines que voulues. Ensuite, il faut remarquer que la procédure est particulièrement adaptée à ce cadre McKean-Vlasov, le support de la mesure discrète* $\hat{\mu}$ *coïncidant avec les points pour lesquelles la solution* $\hat{u}$ *est approchée. Enfin, étant donné l'arbre* $\mathcal{T}$*, il est possible de calculer plusieurs composantes rétrogrades différentes le long de ce même arbre.*

*On conclura cette remarque en soulignant que la complexité exponentielle des arbres de cubature est un problème auquel se sont déjà attelés Litterer et Lyons. Dans [LL12], ils donnent la démarche à effectuer pour réduire la taille du support des mesures* $\hat{\mu}_{T_k}$, $k = 1, \ldots, N$*. L'algorithme obtenu a alors une complexité polynomiale. Malheureusement, la procédure qu'ils proposent étant basée sur une redistribution des points est des poids, les espérances conditionnelles permettant le calcul de la composante rétrograde ne sont plus disponibles. Il serait intéressant de pouvoir adapter cette procédure de manière à contourner ce problème.*

### *2.4* Analyse de l'erreur.

La propriété de Markov permet de ramener l'étude de l'erreur globale de chacun des trois algorithmes en la somme d'erreurs locales (sur chaque nœud de l'arbre). Ensuite, c'est la construction des trois algorithmes et le gel de la loi $(\mathcal{C})$ qui donnent la décomposition des erreurs locales en trois termes : une erreur de schéma, une erreur de cubature et un terme de propagation [24].

Une fois cette décomposition de l'erreur comprise, la preuve repose essentiellement sur l'argument $(\mathcal{C})$, sur des arguments de cubature, des développements de Taylor stochastiques et sur une identification soigneuse des termes de l'expansion et des termes de l'algorithme. Chacun de ces termes est majoré indépendamment et on aboutit à une équation de propagation locale des erreurs. On conclut la preuve grâce à un argument de stabilité.

Cette étude est détaillée en section 2.6 du chapitre 2.

### *2.5* Extensions à un système elliptique avec condition terminale lipschitzienne.

Lorsque la matrice de diffusion est elliptique, on peut étendre les résultats au cas d'une condition terminale lipschitzienne. Dans ce cas, en mettant à nouveau à profit le gel de la loi $(\mathcal{C})$ qui intervient dans la partie McKean-Vlasov du processus, l'extension se fait sur le même procédé que celui expliqué dans le paragraphe pour la cubature.

On recourt donc aux effets régularisants en utilisant des bornes du type de (2.4), l'extension de ces bornes au cas rétrograde étant assurée par le récent travail de Crisan et Delarue [CD11].

Contrairement au cas de la cubature, l'affaiblissement de l'hypothèse d'ellipticité à une hypothèse UFG n'est, ici, pas évidente. En effet, les estimations de gradients de la solution sous des hypothèses UFG ne sont disponibles que pour des coefficients homogènes et la

24. S'ajoute à cela une erreur de prédiction pour le schéma d'ordre 2.

littérature sur l'extension des bornes de gradients au cas inhomogène (voir [CM02]) est assez restreinte. L'astuce de gel ne permettant pas de contourner l'inhomogénéité en temps des coefficients, c'est donc l'ellipticité de la matrice qui garantit l'effet régularisant.

### *2.6* Perspectives.

(1) Affaiblissement de la régularité des coefficients en la loi pour le système elliptique. Lorsque la dépendance des coefficients en la loi est du type (2.8) et que la matrice $V$ est uniformément elliptique, une question naturelle est de savoir comment se comporte l'algorithme lorsque les fonctions $\varphi_i, i = 1, \ldots, d$ sont elles aussi seulement lipschitziennes : un effet régularisant est à nouveau possible.

En annexe B, on montre que les espérances des fonctionnelles de la solution sont différentiable en la condition initiale et satisfont aux bornes de gradients usuelles de type (1.5) en temps court. Les calculs reposent sur une représentation en série parametrix [MS67] de la densité de transition du processus.

A titre d'application, ceci permettrait d'étendre l'algorithme à une classe plus générale de processus, et démontrerait que la vitesse de convergence pour les processus du type (2.1) dont les coefficients sont lipschitziens par rapport à la distance de Wasserstein-1 [25] est d'ordre 1.

Il faut cependant faire attention aux "sens" des deux effets régularisants (2.4) : comme évoqué plus haut, les deux EDP ne sont pas dans la même direction. Il faudrait ainsi choisir une subdivision suffisamment fine aux deux extrémités.

(2) Ordres supérieurs pour le cas d'une dépendance générale des coefficients en la loi. En comparant les ordres (2.7) du Théorème 2.1 et (2.9) du Théorème 2.2, on comprend que la vitesse obtenue dans le cas où la dépendance en la loi est générale n'est pas optimale. Quels que soient $q$ et $m$, cette vitesse est de $N^{-1}$ pour une dépendance générale et $N^{[(m-1)\wedge q]/2}$ pour la dépendance particulière (2.8).

Pour retrouver ces ordres, il faut définir un développement des coefficients dans le cas général. Cela nécessite, avant toutes choses, de définir une différentiation.

Une direction est donnée dans les notes de Cardaliaguet [26] (voir la Section 6 de [Car10]). Il s'agit d'identifier la différentiation d'une fonction le long d'une mesure à la différentiation d'une fonction d'une variable aléatoire appartenant à $L^2(\Omega)$. On peut alors bénéficier de la structure hilbertienne de $L^2(\Omega)$ et définir une différentiation de Fréchet.

(3) Vitesses d'approximation d'algorithmes pour des EDSPR dont la condition terminale n'est pas lipschitzienne. Les bornes de gradients, disponibles pour des conditions moins régulières que Lipschitz, posent naturellement la question de l'affaiblissement de la régularité de la condition terminale. La question n'est pas restreinte au contexte McKean-Vlasov exposé ici. Par exemple, elle a déjà été explorée dans le cas de la composante progressive dans l'article de Talay et Bally [BT96]. Cependant, il semble que la question n'ait pas été posée pour les algorithmes sur la partie rétrograde, du moins pour l'obtention de schémas d'ordre 1 ou plus (*e.g.* [DM06, CM12]).

Récemment, Crisan et Delarue ont étudié les bornes de gradients des solutions d'EDP semi-linéaires sous des conditions UFG. Dans leur travail [CD11],

25. Grâce au Théorème de Monge-Kantorovich.

26. Il semble que l'idée soit celle de Lions, puisque qu'elle a été introduite lors des cours au collège de France.

ils montrent que lorsque la condition terminale est lipschitzienne, on retrouve les mêmes décroissances que celles données par Kusuoka et Stroock [KS84, KS85, KS87]. Cependant, l'étude montre que lorsque la condition terminale est "seulement" bornée, les bornes se dégradent : cette dégradation est due à la structure non-linéaire de l'EDP. Grossièrement, au delà d'un certain ordre de dérivation, c'est la vitesse du produit des bornes sur la dérivée de la solution, issue de la différentiation du terme non-linéaire, qui donne la vitesse de décroissance et non les dérivées successives de la solution. Dans le cas elliptique, un rapide calcul montre que cette vitesse est supérieure à celle du terme linéaire.

Une lecture attentive du travail de Crisan et Delarue permet de comprendre que le cas d'une condition Hölder se déduit comme une interpolation du cas lipschitzien au cas borné. Il serait donc intéressant de savoir quelles vitesses peuvent être envisagées dans ce cadre.

# CHAPITRE 1

## Strong existence and uniqueness for stochastic differential equation with Hölder drift and degenerate noise

### *1.1* Introduction

Let $T > 0$ and $d$ be a positive integer, we consider the following $\mathbb{R}^d \times \mathbb{R}^d$ system for any $t$ in $[0, T]$ :

$$\begin{cases} dX_t^1 = F_1(t, X_t^1, X_t^2)dt + \sigma(t, X_t^1, X_t^2)dW_t, & X_0^1 = x_1, \\ dX_t^2 = F_2(t, X_t^1, X_t^2)dt, & X_0^2 = x_2, \end{cases} \tag{1.1.1}$$

where $(W_t, t \geq 0)$ is a standard $d$-dimensional Brownian motion defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ and $F_1, F_2, \sigma : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{M}_d(\mathbb{R})$ (the set of real $d \times d$ matrices) are measurable functions. The diffusion matrix $a := \sigma\sigma^*$ is supposed to be uniformly elliptic. The notation "$*$" stands for the transpose.

In this Chapter, we investigate the well posedness of (1.1.1) outside the Cauchy-Lipschitz framework. Notably, we are interested in the strong posedness, *i.e* strong existence and uniqueness of a solution. Strong existence means that there exists a process $(X_t^1, X_t^2,\ 0 \leq t \leq T)$ adapted to the filtration generated by the Brownian motion $(W_t,\ 0 \leq t \leq T)$ which satisfies (1.1.1). Strong uniqueness means that, if two processes satisfy this equation with the same initial conditions, their trajectories are almost surely indistinguishable. Here, we show that under a suitable Hölder assumption on the drift coefficients and Lipschitz condition on the diffusion matrix, the strong well-posedness holds for (1.1.1).

It may be a real challenge to prove the existence of a unique solution for a differential system without Lipschitz conditions on the coefficients. For example, in [DL89], DiPerna and Lions showed that under integrability conditions on $b$, $\nabla b$ and $\mathrm{div}(b)$, the integral equation : $Y_t = \int_0^t b(s, Y_s)ds,\ Y_0 = y$ admits a unique solution defined as a regular Lagrangian flow (see [DL89] for the definition of such a solution).

In a stochastic case, the first result in this direction is due to Zvonkin. In [Zvo74], the author showed that the strong well-posedness holds for the one-dimensional system

$$Y_t = \int_0^t b(s, Y_s)ds + W_t,\ Y_0 = y \quad t \in [0, T], \tag{1.1.2}$$

for a measurable function $b$ in $\mathbb{L}^\infty$. Then, Veretennikov [Ver80] generalized this result to the multi-dimensional case and Krylov and Röckner showed in [KR05] the strong well-posedness for $b$ in $\mathbb{L}^p_{\text{loc}}$, $p > d$. There are some extensions of these works and we refer the reader to the paper of Zhang [Zha05] and references therein for a summary of the results. Finally, when $b$ is a measurable and bounded function, Davie showed in [Dav07] that for almost every Brownian path, there exists a unique solution for the system (1.1.2). We emphasize that this result implies the strong uniqueness, but the converse is not true. Indeed, in such a case, there exists an a priori set $\Omega' \subset \Omega$ with $\mathbb{P}(\Omega') = 1$ such that for all $\omega$ in $\Omega'$ the solution of (1.1.2) is unique.

All these results rely on the regularization of differential systems by adding a non-degenerate noise, and we refer to [Fla11] for a partial revue on this subject. In (1.1.1), the noise added is completely degenerate w.r.t the degenerate component $X^2$. This sort of system has also been studied by Veretennikov in [Ver83] but without considering any regularization in the degenerate direction. Indeed, the author showed that strong well-posedness holds when the drift is measurable and bounded and the diffusion matrix is Lipschitz w.r.t the non-degenerate component $X^1$ and when both the drift and the diffusion matrix are twice continuously differentiable functions with bounded derivatives w.r.t the degenerate component.

Their proofs rely on the deep connection between SDEs and PDEs (see [Bas98] or [Fri06] for a partial revue in the elliptic and parabolic cases). The generator associated to the Markov process $Y$ is a linear partial differential operator of second order (usually denoted by $\mathcal{L}$) with the transition density of $Y$ as fundamental solution. As explained by Fedrizzi and Flandoli in [FF11] : "if we have a good theory for the PDE :

$$\frac{\partial}{\partial t}u + \mathcal{L}u = \Phi, \text{ on } [0,T] \quad u_T = \mathbf{0}, \tag{1.1.3}$$

where the source term $\Phi$ has the same regularity as the drift, then, we have the main tools to prove strong uniqueness".

In this Chapter, we show that the noise regularizes, even in the degenerate direction, by means of the random drift. Unfortunately, there is a price to pay to balance the degeneracy of the noise. First, the drift must be at least 2/3-Hölder continuous w.r.t the degenerate component. We do not know how sharp is this critical value, but it is consistent with our approach. Secondly, the drift $F_2$ of the second component must be Lipschitz continuous w.r.t the first component and its derivative in this direction has to be uniformly non degenerate : this allows the drift to regularize.

Our proof also relies on regularization properties of the associated PDE, and the "good theory" is here a "strong theory" : a Lipschitz bound on the solution of (1.1.3) and on its derivative w.r.t the first component. In our case, the generator $\mathcal{L}$ is given by : for all $\psi$ in $C^{1,2,1}([0,T] \times \mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ [27] :

$$\begin{aligned}\mathcal{L}\psi(t,x_1,x_2) &= \frac{1}{2}Tr(a(t,x_1,x_2))D^2_{x_1}\psi(t,x_1,x_2)) + [F_1(t,x_1,x_2)]\cdot[D_{x_1}\psi(t,x_1,x_2)] \\ &\quad + [F_2(t,x_1,x_2)]\cdot[D_{x_2}\psi(t,x_1,x_2)]\,.\end{aligned} \tag{1.1.4}$$

27. *i.e. continuously differentiable w.r.t. the first variable, twice continuously differentiable w.r.t. the second variable and once continuously differentiable w.r.t. the third variable.*

where $Tr(a)$ stands for the trace of the matrix $a$ and "$\cdot$" denotes the standard Euclidean inner product on $\mathbb{R}^d$ and where for any $z$ in $\mathbb{R}^d$, the notation $D_z$ means the derivative w.r.t the variable $z$. Here, the operator is not uniformly parabolic. When the coefficients are smooth and when the Lie algebra generated by the vector fields spans the whole space, such an operator admits a smooth fundamental solution (see [Hö67]) : in this case it is said to be hypoelliptic and the coefficients are said to satisfied a Hörmander condition. The assumption on the uniform non-degeneracy of the derivative of the drift $F_2$ w.r.t $x_1$ can be understood as a sort of weak Hörmander condition.

In our case, the form of the degeneracy is a non-linear generalization of Kolmogorov's degeneracy, in reference to the first work [Kol34] of Kolmogorov in this direction. Degenerate operators of this form have been studied by many authors see e.g. the work of Di Francesco and Polidoro [DFP06], and Delarue and Menozzi [DM10]. We also emphasize that, in [Men11], Menozzi deduced from the regularization property exhibited in ([DM10]) the well weak posedness of a generalization of (1.1.1). Nevertheless, to the best of our knowledge, there does not exist a strong theory, in the sense defined above, for the PDE (1.1.3) when $\mathcal{L}$ is defined by (1.1.4). We investigate it by using the so called parametrix approach (see [Fri64] for partial revue in the elliptic setting).

In comparison with the works of Veretennikov [Ver80, Ver83], Krylov and Röckner [KR05], and Flandoli and Fedrizzi [FF11], asking for $F_1$ to be only in $\mathbb{L}^p$, $p > d$ might appear as the right framework. Since the parametrix is a perturbation method and we are interested in $\mathbb{L}^\infty$ estimates, we suppose the drift $F_1$ to be Hölder continuous w.r.t $x_1$.

### *1.1.1* Organization of this Chapter

Subsection 1.1.2 states useful notations, detailed assumptions and the main result of this Chapter : strong existence and uniqueness for (1.1.1). In Subsection 1.1.3, we provide the strategy to prove this result, which includes the regularization properties of the associated PDE. Finally, our main result is proved in Subsection 1.1.4. The remainder of this Chapter is dedicated to the proof of the regularization properties of the associated PDE.

The strategy is exposed in Subsection 1.1.5 : it is based on a smooth approximation of the coefficients and the parametrix. Existence and uniqueness of a smooth solution for the PDE with smooth coefficients is given in Subsection 1.1.6.

Section 1.2 explains the proof of the regularization properties in a simple case and allows to understand our assumptions and how the proof in the general case can be achieved. Section 1.3 defines the mathematical tools and the proof of the regularization properties of the PDE is provided in Section 1.4. This is the technical part of this Chapter.

### *1.1.2* Main Result

**Notations.** In order to simplify the notations, we adopt the following convention : $x, y, z, \xi$, etc.. denote the $2d$−dimensional real variables $(x_1, x_2), (y_1, y_2), (z_1, z_2), (\xi_1, \xi_2)$, etc.. Consequently, each component of the $d$-dimensional variables $x_k$, $k = 1, 2$ is denoted by $x_{kl}$, $l = 1, \cdots, d$. We denote by $g(t, X_t)$ any function $g(t, X_t^1, X_t^2)$ from $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^N$, $N \in \mathbb{N}$. Here, $X_t = (X_t^1, X_t^2)$ and then $F(t, X_t)$ is the $\mathbb{R}^{2d}$ valued function $(F_1(t, X_t^1, X_t^2), F_2(t, X_t^1, X_t^2))^*$. We rewrite the system (1.1.1) in a shortened form :

$$dX_t = F(t, X_t)dt + B\sigma(t, X_t)dW_t, \tag{1.1.5}$$

where $B$ is the $2d \times d$ matrix : $B = (\mathrm{Id}, 0_{\mathbb{R}^d \times \mathbb{R}^d})^*$. "Id" stands for the identity matrix of $\mathcal{M}_d(\mathbb{R})$, the set of real $d \times d$ matrices. When necessary, we write $(X_s^{t,x})_{t \leq s \leq T}$ the process in (1.1.1) which starts from $x$ at time $t$, i.e $X_t^{t,x} = x$.

We denote by $C^{1,2,1}$ the space of functions that are continuously differentiable w.r.t. the first variable, twice continuously differentiable w.r.t. the second variable and and once continuously differentiable w.r.t. the third variable.

We denote by $\mathrm{GL}_d(\mathbb{R})$ the set of $d \times d$ invertible matrices with real coefficients and by $\phi$ a measurable function from $[0,T] \times \mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^2$. Each one-dimensional component of this function is denoted by $\phi_i$, $i = 1, 2$ and plays the role of one coordinate of $F_i$. Hence, $\phi_i$ satisfies the same regularity as $F_i$ given latter. We recall that $a$ denotes the square of the diffusion matrix $\sigma$ : $a := \sigma\sigma^*$. Subsequently, we denote by $c$, $C$, $C'$, $C''$ or $C'''$ a positive constant, depending only on known parameters in **(H)**, given just below, and may change from line to line and from an equation to another.

We recall that the canonical Euclidean inner product on $\mathbb{R}^d$ is denoted by "$\cdot$" and the notation $D_z$ means the derivative w.r.t the variable $z$. Hence, for all integer $n$, $D_z^n$ is the $n^{\text{th}}$ derivative w.r.t $z$ and for all integer $m$ the $n \times m$ cross differentiations w.r.t $z$, $y$ are denoted by $D_z^n D_y^m$. Furthermore, the partial derivative $\partial/\partial_t$ is denoted by $\partial_t$.

**Hypotheses.** ***(H)***. *We say that assumptions* ***(H)*** *hold if the coefficients satisfy :*

***(H1)*** *: Regularity of the coefficients. There exist* $0 < \beta_i^j < 1$, $1 \leq i, j \leq 2$ *and three positive constants* $C_1, C_2, C_\sigma$ *such that for all* $(t, x_1, x_2)$ *and* $(t, y_1, y_2)$ *in* $[0,T] \times \mathbb{R}^d \times \mathbb{R}^d$,

$$\begin{aligned}
|F_1(t,x_1,x_2) - F_1(t,y_1,y_2)| &\leq C_1(|x_1 - y_1|^{\beta_1^1} + |x_2 - y_2|^{\beta_1^2}) \\
|F_2(t,x_1,x_2) - F_2(t,y_1,y_2)| &\leq C_2(|x_1 - y_1| + |x_2 - y_2|^{\beta_2^2}) \\
|\sigma(t,x_1,x_2) - \sigma(t,y_1,y_2)| &\leq C_\sigma(|x_1 - y_1| + |x_2 - y_2|).
\end{aligned}$$

*Moreover, the coefficients are supposed to be continuous w.r.t the time and* $\beta_i^2 > 2/3,\ i = 1, 2$. *Thereafter, we set* $\beta_2^1 = 1$ *for notational convenience.*

***(H2)*** *: Uniform ellipticity of* $\sigma\sigma^*$. *The function* $\sigma\sigma^*$ *also satisfies the uniform ellipticity hypothesis :*

$$\exists \Lambda > 1 \text{ such that } \forall \zeta \in \mathbb{R}^{2d}, \quad \Lambda^{-1}|\zeta|^2 \leq [\sigma\sigma^*(t,x_1,x_2)\zeta] \cdot \zeta \leq \Lambda|\zeta|^2,$$

*for all* $(t, x_1, x_2) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^d$.

***(H3-a)*** *: Differentiability and regularity of* $x_1 \mapsto F_2(., x_1, .)$. *For all* $(t, x_2) \in [0,T] \times \mathbb{R}^d$, *the function* $F_2(t, ., x_2) :\ x_1 \mapsto F_2(t, x_1, x_2)$ *is continuously differentiable and there exist* $0 < \alpha^1 < 1$ *and a positive constant* $\bar{C}_2$ *such that, for all* $(t, x_2)$ *in* $[0,T] \times \mathbb{R}^d$ *and* $x_1, y_1$ *in* $\mathbb{R}^d$,

$$|D_{x_1}F_2(t,x_1,x_2) - D_{x_1}F_2(t,y_1,x_2)| \leq \bar{C}_2|x_1 - y_1|^{\alpha^1}.$$

***(H3-b)*** *: Non degeneracy of* $(D_{x_1}F_2)(D_{x_1}F_2)^*$. *There exists a closed convex subset* $\mathcal{E} \subset \mathrm{GL}_d(\mathbb{R})$ *such that for all* $t$ *in* $[0,T]$ *and* $(x_1, x_2)$ *in* $\mathbb{R}^{2d}$ *the matrix* $D_{x_1}F_2(t, x_1, x_2)$ *belongs to* $\mathcal{E}$. *We emphasize that this implies that*

$$\exists \bar{\Lambda} > 1 \text{ such that } \forall \zeta \in \mathbb{R}^{2d}, \quad \bar{\Lambda}^{-1}|\zeta|^2 \leq [(D_{x_1}F_2)(D_{x_1}F_2)^*(t,x_1,x_2)\zeta] \cdot \zeta \leq \bar{\Lambda}|\zeta|^2,$$

*for all* $(t, x_1, x_2) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^d$.

**Remark.** *Consequently, the sentence "known parameters in **(H)**" refers to the parameters belonging to these assumptions.*

*The reason for the existence of the critical value* $2/3$ *for the Hölder regularity of the drift in **(H1)** and the particular "convexity" assumption **(H3-b)** are discussed in Section 1.2.*

The following Theorem is the main result of this Chapter and regards the strong well-posedness of the system (1.1.1).

**Theorem 1.1.1.** *Let* $T > 0$ *and suppose that assumptions **(H)** hold true. Then, strong existence and uniqueness hold for* (1.1.1).

### *1.1.3* Strategy of proof

Let us expose the basic arguments to prove Theorem 1.1.1. Existence of a weak solution follows from a compactness argument, see e.g. [SV79]. Then, if the strong uniqueness holds, the strong existence follows. The main issue consists in proving the strong uniqueness. As we already mentioned, the strategy relies on regularization properties of the linear system of PDEs :

$$\begin{cases} \partial_t u_i(t,x) + \mathcal{L} u_i(t,x) = F_i(t,x), & \text{for } (t,x) \in [0,T] \times \mathbb{R}^{2d}, \\ u_i(T,x) = 0_{\mathbb{R}^d}, & i = 1,2. \end{cases} \tag{1.1.6}$$

This works as follows : suppose that there exists a unique $C^{1,2,1}([0,T] \times \mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ solution $u = (u^1, u^2)^*$ of this system such that $u$ and $D_x u$ are $C_T$ Lipschitz continuous, where $C_T$ is small as $T$ is small. Thanks to Itô's formula, for all $t$ in $[0,T]$ :

$$\int_0^t F(s, X_s) ds = u(t, X_t) - \int_0^t D_x u(s, X_s) B \sigma(s, X_s) dB_s.$$

Then, $\int_0^t F(s, X_s) ds$ is Lipschitz continuous w.r.t $X$, with Lipschitz constant $C_T$. The uniqueness can be proved for small $T$ by a circular argument and the result can be deduced by iterating this strategy on sufficiently small partitions of the interval.

The main issue here is that the PDE (1.1.6) does not admit a $C^{1,2,1}([0,T] \times \mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ solution. Nevertheless, we do not need to obtain the existence of a regular solution but only the existence of Lipschitz bounds for $u$ and $D_x u$ depending only on known parameters in **(H)**. Therefore, we investigate these bounds in a smooth setting. Indeed, thanks to assumptions **(H)**, there exists a sequence of smooth mollified coefficients $(a^n, F_1^n, F_2^n)_{n \geq 1}$ with bounded derivatives of any order such that :

$$(a^n, F_1^n, F_2^n) \underset{n \to +\infty}{\longrightarrow} (a, F_1, F_2), \tag{1.1.7}$$

uniformly on $[0,T] \times \mathbb{R}^d \times \mathbb{R}^d$ and such that $(a^n, F_1^n, F_2^n)$ satisfy **(H)** uniformly (in $n$). More details on the regularization procedure are given in Subsection 1.1.5 below. Let us denote by $\mathcal{L}^n$ the regularized version of $\mathcal{L}$ (that is the version of $\mathcal{L}$ with mollified coefficients), one has :

**Lemma 1.1.2.** *Let* $n$ *in* $\mathbb{N}^*$. *The PDE,*

$$\partial_t u_i^n(t,x) + \mathcal{L}^n u_i^n(t,x) = F_i^n(t,x), \quad \textit{for } (t,x) \in [0,T] \times \mathbb{R}^{2d} \quad u_i^n(T,x) = 0_{\mathbb{R}^d}, \quad i = 1,2,$$

*admits a unique solution* $u^n = (u_1^n, u_2^n)^*$, *which is infinitely differentiable with bounded derivatives of any order.*

Moreover the solutions $u^n,\ n \geq 1$ satisfy :

**Proposition 1.1.3.** *For $T$ small enough, there exist a positive constant $C$ and a positive number $\delta_{1.1.3}$ depending only on $T$ and known parameters in **(H)** and not on $n$ such that :*

$$||D_{x_1} u^n||_\infty + ||D_{x_2} u^n||_\infty + ||D^2_{x_1} u^n||_\infty + ||D_{x_1} D_{x_2} u^n||_\infty \leq C T^{\delta_{1.1.3}}.$$

We emphasize that the estimates on the solutions $u^n,\ n \geq 1$ are obtained uniformly in $n$ (that is independently of the procedure of regularization) and we do not have to solve the limit PDE problem. Besides, the terminal condition $u^n(T, .) = 0$ is very important : it guarantees that the solution and its derivatives vanish at time $T$. Hence, it allows to control the Lipschitz constant of $u^n$ by a constant small as $T$ is small.

### *1.1.4* Proof of Theorem 1.1.1

As we discussed, we only have to prove the strong uniqueness, since the weak existence holds. Thereafter, we denote by "**1**" the $2d \times 2d$ matrix :

$$\begin{pmatrix} \mathrm{Id} & 0_{\mathbb{R}^d \times \mathbb{R}^d} \\ 0_{\mathbb{R}^d \times \mathbb{R}^d} & 0_{\mathbb{R}^d \times \mathbb{R}^d} \end{pmatrix}. \tag{1.1.8}$$

Let $(X_t, t \geq 0)$ and $(Y_t, t \geq 0)$ be two solutions of (1.1.1) with the same initial condition $x$ in $\mathbb{R}^{2d}$. Let $u^n$ be the solution of the linear system of PDEs (1.1.6). By using Lemma 1.1.2 we can apply Itô's formula for $u^n(t, X_t) - X_t$ and we obtain :

$$\begin{aligned} u^n(t, X_t) - X_t &= \int_0^t \left[\partial_t u^n + \mathcal{L} u^n\right](s, X_s) ds - \int_0^t F(s, X_s) ds + u^n(0, x) - x \\ &\quad + \int_0^t \left[D_x u^n - \mathbf{1}\right] B\sigma(s, X_s) dW_s. \end{aligned}$$

In order to use the fact that $\partial_t u^n + \mathcal{L}^n u^n = F^n$, we rewrite :

$$\begin{aligned} u^n(t, X_t) - X_t &= \int_0^t \left[\partial_t u^n + \mathcal{L}^n u^n\right](s, X_s) ds + \int_0^t \left(\mathcal{L} - \mathcal{L}^n\right) u^n(s, X_s) ds \\ &\quad - \int_0^t F(s, X_s) ds + u^n(0, x) - x \\ &\quad + \int_0^t \left[D_x u^n - \mathbf{1}\right] B\sigma(s, X_s) dW_s, \end{aligned}$$

and then,

$$\begin{aligned} u^n(t, X_t) - X_t &= \int_0^t \left(\mathcal{L} - \mathcal{L}^n\right) u^n(s, X_s) ds + \int_0^t (F^n(s, X_s) - F(s, X_s)) ds + u^n(0, x) - x \\ &\quad + \int_0^t \left[D_x u^n - \mathbf{1}\right] B\sigma(s, X_s) dW_s, \end{aligned}$$

By the same argument, we obtain :

$$\begin{aligned} u^n(t, Y_t) - Y_t &= \int_0^t \left(\mathcal{L} - \mathcal{L}^n\right) u^n(s, Y_s) ds + \int_0^t (F^n(s, Y_s) - F(s, Y_s)) ds + u^n(0, x) - x \\ &\quad + \int_0^t \left[D_x u^n - \mathbf{1}\right] B\sigma(s, Y_s) dW_s. \end{aligned}$$

By taking the expectation of the supremum over $t$ of the square norm of the difference of the two equalities above, we get :

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}|X_t-Y_t|^2\right] \;\leq\; & C\mathbb{E}\left[\sup_{t\in[0,T]}|u^n(t,X_t)-u^n(t,Y_t)|^2\right] \\
& +C\mathbb{E}\left[\int_0^T \left|\left[D_x u^n B\right](s,X_s)-\left[D_x u^n B\right](s,Y_s)\right|^2 |\sigma(s,Y_s)|^2\, ds\right] \\
& +C\mathbb{E}\left[\int_0^T (\|D_x u^n B\|_\infty+1)\,|[\sigma(s,Y_s)-\sigma(s,X_s)]|^2\, ds\right] \\
& +C\mathcal{R}(n,T),
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{R}(n,T) \;=\; & CT\left(\mathbb{E}\left[\sup_{t\in[0,T]}|F^{(n)}(t,Y_t)-F(t,Y_t)|^2\right]+\mathbb{E}\left[\sup_{t\in[0,T]}|(\mathcal{L}^n-\mathcal{L})u^n(t,Y_t)|^2\right]\right) \\
& +CT\left(\mathbb{E}\left[\sup_{t\in[0,T]}|F^{(n)}(t,X_t)-F(t,X_t)|^2\right]+\mathbb{E}\left[\sup_{t\in[0,T]}|(\mathcal{L}^n-\mathcal{L})u^n(t,X_t)|^2\right]\right).
\end{aligned}$$

First, note that from (1.1.7), for both $Y_t$ and $X_t$, we have :

$$\mathbb{E}\left[\sup_{t\in[0,T]}|F^n(t,X_t)-F(t,X_t)|^2\right]+\mathbb{E}\left[\sup_{t\in[0,T]}|(\mathcal{L}^n-\mathcal{L})u^n(t,X_t)|^2\right]\to 0,\ \text{ as } n\to\infty,$$

so that $\mathcal{R}(n,T)\to 0$ when $n\to\infty$. Secondly, we know from Proposition 1.1.3, that for $T$ small enough and for all $t\in[0,T]$, the functions $u^n(t,.,.)$ and $D_{x_1}u^n(t,.,.)$ are Lipschitz continuous, with a Lipschitz constant independent of $n$. Since $D_x u^n B=(D_{x_1}u^n, 0_{\mathbb{R}^d\times\mathbb{R}^d})$, by letting $n\to\infty$ and using the two arguments above, we deduce that for $T$ small enough :

$$\mathbb{E}\left[\sup_{t\in[0,T]}|X_t-Y_t|^2\right]\leq C(T)\left\{\mathbb{E}\left[\sup_{t\in[0,T]}|X_t-Y_t|^2\right]+\mathbb{E}\left[\int_0^T |X_s-Y_s|^2 ds\right]\right\},$$

where $C(T)\to 0$ when $T\to 0$. Then, the strong uniqueness holds for $T$ small enough. By iterating this computation, the same result holds on any finite intervals and so on $[0,\infty)$. □

### *1.1.5* Strategy of proof of Lemmas 1.1.2 and Proposition 1.1.3

Firstly, we emphasize that each coordinate of the vectorial solution $u_i$ of the decoupled linear PDE (1.1.6) can be described by the PDE

$$\partial_t u_i(t,x)+\mathcal{L}u_i(t,x)=\phi_i(t,x),\quad \text{for } (t,x)\in[0,T]\times\mathbb{R}^{2d},\quad u_i(T,x)=0, \tag{1.1.9}$$

where $\phi_i:\mathbb{R}^{2d}\to\mathbb{R}$ satisfies the same regularity assumptions as $F_i$ given in **(H1)**. Therefore, we only have to prove Lemma 1.1.2 and Proposition 1.1.3 for (1.1.9) instead of (1.1.6).

Then, we do not solve the limit PDE problem (1.1.6). The investigations are done with mollified coefficients $a^n, F_1^n, F_2^n, \phi_1^n, \phi_2^n,\ n \geq 1$ [28].

Let us detail how the smooth approximation of the coefficients $a, F_1, F_2, \phi_1, \phi_2$ works. For all $n$ in $\mathbb{N}^*$, we set :

$$F_2^n(t,x) = \int F_2(t-s, x-y)\varphi_1^n(y)\varphi_2^n(s)dyds,$$

where $\varphi_1^n(.) = c_1 n^{2d}\varphi(n|.|)$ and $\varphi_2^n(.) = c_2 n\varphi(n|.|)$ for $c_1$, $c_2$ two constants of normalization and for a smooth function $\varphi$ with support in the unit ball. For example $\varphi : z \in \mathbb{R} \mapsto \exp\left(-\frac{1}{1-z^2}\right)\mathbf{1}_{]-1;1[}(z)$ and $|.| = ||.||_\infty$. Defined $(F_1^n)_{n\geq 1}, (\phi_1^n)_{n\geq 1}, (\phi_2^n)_{n\geq 1}$ and $(a^n)_{n\geq 1}$ with the same procedure, it is then clear that for every $n$ in $\mathbb{N}^*$ the mollified coefficients $a^n, F_1^n, F_2^n, \phi_1^n, \phi_2^n$ are smooth with bounded derivatives of any order and such that $(a^n, F_1^n, F_2^n, \phi_1^n, \phi_2^n)_{n\geq 1}$ converges uniformly on $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$ to $(a, F_1, F_2, \phi_1, \phi_2)$. Moreover, they satisfy the same assumptions as $(a, F_1, F_2, \phi_1, \phi_2)$ uniformly in $n$. Let us just check the non-degeneracy assumption **(H3-a)** on $D_{x_1}F_2^n$. For all $\delta > 0$, one can find a sequence of rectangles $(R_k)_{1\leq k\leq N(\delta)}$ having sides of length less than $\delta$ and a family of points $\{(s_k, y_k) \in R_k,\ 1 \leq k \leq N(\delta)\}$ such that, for all $(t,x) \in [0,T]\times\mathbb{R}^{2d}$ :

$$D_{x_1}F_2^n(t,x) = \lim_{\delta\to 0}\sum_{k=1}^{N(\delta)} D_{x_1}F_2(t-s_k, x-y_k)\int_{R_k}\varphi_1^n(y)\varphi_2^n(s)dyds.$$

Since $D_{x_1}F_2$ belongs to the closed convex subset $\mathcal{E}$, it is clear that $D_{x_1}F_2^n$ belongs to $\mathcal{E}$.

Hence, the estimates in Proposition 1.1.3 are obtained for mollified coefficients but depend only on known parameters in **(H)**. Consequently, we forget the superscript "$n$" which arises from the mollifying procedure, and we further assume that :

**Hypotheses.** ***(HR)*** *: We say that assumptions* ***(HR)*** *hold if : Assumptions* ***(H)*** *hold true and $F_1, F_2, \phi_1, \phi_2, a$ are infinitely differentiable functions with bounded derivatives of any order.*

The existence of a smooth solution under **(HR)** is established in Subsection 1.1.6 below. Then, the estimates on this solution are obtained by using the parametrix method (see [Fri64] for a revue in the elliptic setting).

### *1.1.6* Proof of Lemma 1.1.2

**Strategy.** We show that under **(HR)**, there exists a unique solution $u$ of the linear system of PDEs (1.1.9) which is infinitely differentiable with bounded derivatives of any order. Existence and regularity are proved by adopting a viscosity solution approach, and uniqueness by using the Feynman-Kac representation. This proof is decomposed as follows :

We propose a candidate $u_i$ as :

28. here, $(\phi_1^n, \phi_2^n)_{n\geq 1}$ denotes the sequence of mollified coefficients $(\phi_1, \phi_2)$ : they are infinitely differentiable with bounded derivatives of any order and they satisfy the same hypotheses as the $(\phi_1, \phi_2)$ uniformly in $n$.

$$u_i(t,x_1,x_2) = \mathbb{E}\left[\int_t^T \phi_i(s, X_s^{1,t,x}, X_s^{2,t,x}) ds\right], \tag{1.1.10}$$

where the process $(X^1, X^2)$ satisfies EDS (1.1.1).
(i) We show that $u_i$ is a viscosity solution of the linear systems (1.1.9).
(ii) We show that for all $t$ in $[0,T]$, the function $u_i(t,.,.) : (x_1,x_2) \in \mathbb{R}^{2d} \mapsto u_i(t,x_1,x_2)$ is infinitely differentiable with bounded derivatives, thanks to the property of differentiation of the mapping $(X^{1,\cdot}, X^{2,\cdot}) : x \in \mathbb{R}^{2d} \mapsto (X^{1,t,x}, X^{2,t,x})$.
(iii) From the Markov property (that holds under **(HR)**), we deduce that, for all $(x_1,x_2)$ in $\mathbb{R}^{2d}$, the function $u_i(.,x_1,x_2) : \ t \in [0,T] \mapsto u_i(t,x_1,x_2)$ is infinitely differentiable with bounded derivatives.
(iv) Since a smooth viscosity solution is a classical solution, we obtain the existence and regularity of a classical solution of (1.1.9).
(v) We conclude the proof of uniqueness by using the Feynman-Kac representation of the solution.

**Proof.** Let $i \in \{1,2\}$ and consider the function $u_i$ defined in (1.1.10) :

(i) Under **(HR)** we deduce from [Kun82] and Theorem 70 of [Pro04] that there exists a unique strong solution $X = (X^1, X^2)$ of the stochastic system and an a.s. continuous version of this process $(X_s^{1,t,x}, X_s^{2,t,x})_{t\le s\le T}$. From the regularity of $\phi_i$, we deduce that $u_i$ is continuous. By using Theorem 5.2 p 190 of [YZ99], we conclude that $u_i$ is a sub and super viscosity solution.

(ii) Thanks to [Kun82], we know that for all $t$ in $[0,T]$, for all $s$ in $[t,T]$, the mapping $X_s^{t,\cdot} : x \mapsto X_s^{t,x}$ is a.s. infinitely differentiable and, for all $k$ in $\mathbb{N}^*$, for all $(i_1, \cdots, i_k) \in \{1,2\}^k$ the tangent process $(D^k_{x_{i_1},\cdots x_{i_k}} X_s^{t,x})_{t\le s\le T}$ satisfies :

$$\mathbb{E}\left[\sup_{s\in[t,T]} \left|D^k_{x_{i_1},\cdots x_{i_k}} X_s^{t,x}\right|\right] \le K.$$

Since $\phi_i$ is Lipschitz continuous, it satisfies the domination property :

$$|\phi_i(X_T^{t,x}) - \phi_i(X_T^{t,z})| \le K'|x-z|,$$

where $K'$ is a random constant with finite moments of all order according to the Kolmogorov Theorem. Then, one can apply the Lebesgue differentiation Theorem on the right hand side of (1.1.10) and, for all $t$ in $[0,T]$, we deduce that $D_{x_j} u_i(t,x_1,x_2)$, $j=1,2$ exist and satisfy

$$D_{x_j} u_i(t,x_1,x_2) = \mathbb{E}\left[\sum_{l=1}^2 D_{x_l}\phi_i(X_T^{t,x}) D_{x_j} X_T^{l,t,x}\right]. \tag{1.1.11}$$

By iterating this argument, we obtain that $u_i$ is infinitely differentiable w.r.t the space variables $(x_1,x_2)$ and its derivatives are bounded.

(iii) On a first hand, we know that $(X_s^{1,t,x}, X_s^{2,t,x})$ is continuous w.r.t $t$ (see Lemma 4.6.1 of [Kun82]). So that for all $x$ in $\mathbb{R}^{2d}$, the function $u_i(.,x_1,x_2) : \ t \in [0,T] \mapsto u_i(t,x_1,x_2)$ is continuous. By using the Markov property we deduce that, for all $0 < h < t$ :

$$u_i(t-h,x_1,x_2) = \mathbb{E}\left[u_i(t,X_t^{1,t-h,x},X_t^{2,t-h,x})\right] - \mathbb{E}\left[\int_{t-h}^t \phi_i(s,X_s^{1,t-h,x},X_s^{2,t-h,x})ds\right]. \tag{1.1.12}$$

On the other hand, by applying Itô's formula on $\mathbb{R}^d \times \mathbb{R}^d$ for $u_i(t,X_t^{1,t-h,x},X_t^{2,t-h,x})$ and taking the expectation we have :

$$\begin{aligned}
&\mathbb{E}\left[u_i(t,X_t^{1,t-h,x},X_t^{2,t-h,x})\right] \\
&= u_i(t,x_1,x_2) + \mathbb{E}\left[\int_{t-h}^t \mathcal{L}u_i(t,X_r^{1,t-h,x},X_r^{2,t-h,x})dr\right] \\
&+\mathbb{E}\left[\int_{t-h}^t D_x u_i(t,X_r^{1,t-h,x},X_r^{2,t-h,x})B\sigma(r,X_r^{1,t-h,x},X_r^{2,t-h,x})dW_r\right],
\end{aligned} \tag{1.1.13}$$

where the last term in the right hand side is equal to 0.

By summing (1.1.12) and (1.1.13), we obtain :

$$\begin{aligned}
&\frac{u_i(t-h,x_1,x_2)-u_i(t,x_1,x_2)}{h} \\
&= \frac{1}{h}\mathbb{E}\left[\int_{t-h}^t \mathcal{L}u_i(t,X_r^{1,t-h,x},X_r^{2,t-h,x}) - \phi_i(r,X_r^{1,t-h,x},X_r^{2,t-h,x})dr\right].
\end{aligned}$$

The continuity of $u_i$ w.r.t $t$ and letting $h$ tends to 0 give :

$$-\frac{\partial u_i}{\partial t}(t,x_1,x_2) = \mathcal{L}u_i(t,x_1,x_2) - \phi_i(t,x_1,x_2).$$

(iv) Then, by iterating this argument and using the boundedness of the tangent process at every order we deduce that the function $u_i$ defined in (1.1.10) is infinitely differentiable with bounded derivatives of any order. Since a smooth viscosity solution is a classical solution, this concludes the existence part.

(v) Let $v_i$ be such a solution, by applying Itô's formula for $v_i(T,X_T^{1,t,x},X_T^{2,t,x})$, where $(X^{1,t,x},X^{2,t,x})$ is a solution of the SDE (1.1.1) such that $(X_t^{1,t,x},X_t^{2,t,x}) = (x_1,x_2)$ a.s. and taking the expectation we have :

$$v_i(t,x_1,x_2) = \mathbb{E}\left[\int_t^T \phi_i(s,X_s^{1,t,x},X_s^{2,t,x})ds\right],$$

then, $v_i = u_i$ and the uniqueness follows from uniqueness in law for (1.1.1) under **(HR)**. This concludes the proof of Lemma 1.1.2. □

## *1.2* The linear and Brownian heuristic

This section introduces the main issue when proving Proposition 1.1.3 in a simple case. Furthermore, it allows to understand the role of some of the assumptions in **(H)** and to present in a simple form the effects of the degeneracy. By "simple", we mean that the following assumptions hold :

**Hypotheses.** ***(HL)*** *We say that hypotheses* ***(HL)*** *hold if* ***(H)*** *and* ***(HR)*** *hold with :* $F_1 \equiv 0_{\mathbb{R}^d}$, $\sigma \equiv \mathrm{Id}$, *for all* $(t < s, x_1, x_2) \in [0,T]^2 \times \mathbb{R}^{2d}$, $F_2(s,x_1,x_2) = \bar{F}_2(x_2) + \Gamma_s x_1$. *This implies that for all $s$ in*$[0,T]$, $\Gamma_s$ *belongs to the convex subset* $\mathcal{E}$ *of* $\mathrm{GL_d}(\mathbb{R})$.

The SDE (1.1.1) becomes :

$$\left\{\begin{array}{ll} dX_s^1 = dW_s, & X_t^1 = x_1, \\ dX_s^2 = (\bar{F}_2(X_s^2) + \Gamma_s X_s^1)ds, & X_t^2 = x_2, \end{array}\right. \tag{1.2.1}$$

for all $t < s$ in $[0,T]^2$, $x$ in $\mathbb{R}^d$, and under **(HL)** this system admits a unique strong solution $X$. We recall that the associated PDE is :

$$\left\{\begin{array}{l} \partial_t u_i(t,x) + \mathcal{L}u_i(t,x) = \phi_i(t,x), \quad \text{for } (t,x) \in [0,T] \times \mathbb{R}^{2d} \\ u_i(T,x) = 0, \quad i = 1,2 \end{array}\right. \tag{1.2.2}$$

We first present how the parametrix approach works for obtaining a representation of the solution of (1.2.2). Then, we estimate this solution in term on known parameters in **(H)** when the assumption $\beta_2^2 > 2/3$ is satisfied.

### *1.2.1* The frozen system and parametrix

As we discussed, we study the solution of the PDE (1.2.2) by mean of parametrix approach. Let us first introduce the main idea behind this method. It is based on the following observation : In small time, the generator of the solution of an SDE with smooth and variable coefficients and the generator of the solution of the same SDE with fixed (frozen) coefficients are "closed". In the elliptic setting, the frozen generator is the Laplacian. In other words, this consists in considering the variable generator as a perturbation of the Laplacian. The choice of the freezing point plays a central role : it has to be of the order of the typical trajectories of the frozen solution.

Below, we first show how to obtain the frozen system, next, we study the properties the properties of this system and deduce a representation of the solution of (1.2.2).

**Kolmogorov's example.** Let $d = 1$, in [Kol34] Kolmogorov showed that the solution of : $dY_s^1 = dW_s,\ dY_s^2 = \alpha Y_s^1 ds$, $(\alpha \neq 0)$, with initial condition $(Y_0^1, Y_0^2) = (x_1, x_2) \in \mathbb{R}^2$, admits a density. Notably, this density is Gaussian and given by, for all $s$ in $(0,T]$ and $(y_1, y_2) \in \mathbb{R}^2$ :

$$p(0, x_1, x_2; s, y_1, y_2) = \frac{\sqrt{3}}{\alpha \pi s^2} \exp\left(-\frac{1}{2}\left|K_s^{-1/2}(y_1 - x_1, y_2 - x_2 - s\alpha x_1)^*\right|^2\right), \tag{1.2.3}$$

with the following covariance matrix $K_s$ :

$$K_s := \begin{pmatrix} s & (1/2)\alpha s^2 \\ (1/2)\alpha s^2 & (1/3)\alpha^2 s^3 \end{pmatrix}. \tag{1.2.4}$$

This example illustrates the behaviour of the system in small time : it is not diffusive. The first coordinate oscillates with fluctuations of order $1/2$, while the second one oscillates with fluctuations of order $3/2$. As a direct consequence, the transport of the initial condition of the first coordinate has a key role in the second one.

**The frozen system.** In the parametrix method, the choice of the freezing points has to be done carefully. As the Kolmogorov's example suggested, we have to to keep track of the transport of the initial condition. Then, we freeze the system (1.2.1) along the curve $\theta_{t,s} = \left(\theta_{t,s}^1, \theta_{t,s}^2\right)^*$, $s$ in $[t,T]$ that solves the ODE :

$$\frac{d}{ds}\theta_{t,s} = \left(0_{\mathbb{R}^d},\ \bar{F}_2(\theta_{t,s}^2(\xi)) + \Gamma_s \theta_{t,s}^1(\xi)\right)^*,\ \theta_{t,t}(\xi) = \xi, \tag{1.2.5}$$

for all $\xi$ in $\mathbb{R}^{2d}$. This curve can be written as :

$$\theta^1_{t,s}(\xi) = \xi^1, \quad \theta^2_{t,s}(\xi) = \xi^2 + \int_t^s \left[\Gamma_r \xi^1 + \bar{F}_2(\theta^2_{t,r}(\xi))\right] dr. \tag{1.2.6}$$

So, the frozen system is :

$$\begin{cases} d\bar{X}^1_s = dW_s, & \bar{X}^1_t = x_1, \\ d\bar{X}^2_s = \left(\bar{F}_2(\theta^2_{t,s}(\xi)) + \Gamma_s \bar{X}^1_s\right) ds, & \bar{X}^2_t = x_2, \end{cases} \tag{1.2.7}$$

for all $s$ in $(t,T]$. This is our candidate to approximate (1.2.1). Obviously, in order to reproduce the typical trajectories of the frozen process, the variable $\xi$ in (1.2.5) will be chosen to be the initial condition $x$ of teh solution of the SDE (1.2.7).

### *1.2.2* Existence of a density for the frozen system

In this case, the crucial point is the specific form of the covariance matrix $\bar{\Sigma}_{t,\cdot}$ of $\bar{X}_\cdot$. For any $s$ in $(t,T]$, standard computations show that :

$$\bar{\Sigma}_{t,s} = \begin{pmatrix} (s-t) & \int_t^s \int_t^r \Gamma_u du dr \\ \int_t^s \int_t^r \Gamma_u du dr & \int_t^s \left(\int_t^r \Gamma_u du\right)\left(\int_t^r \Gamma_u du\right)^* dr \end{pmatrix}. \tag{1.2.8}$$

Therefore, the existence of a transition density of $\bar{X}^x_s$ results from the non-degeneracy of $\bar{\Sigma}_{t,s}$. In their work [DM10] (see Proposition 3.1), Delarue and Menozzi show that a sufficient condition is given by $\Gamma_r \in \mathrm{GL}_d(\mathbb{R})$ for a.e. $r \in [t,s]$. This gives :

**Lemma 1.2.1.** *Suppose that assumptions **(HL)** hold true, then, the solution of (1.2.7) admits a transition density $\bar{q}$ given by, for all $s$ in $(t,T]$ :*

$$\bar{q}(t,x_1,x_2;s,y_1,y_2) = \frac{1}{(2\pi)^{d/2}} \det(\bar{\Sigma}_{t,s})^{-1/2} \exp\left(-|\bar{\Sigma}^{-1/2}_{t,s}(y_1 - x_1, y_2 - m^{2,\xi}_{t,s}(x))^*|^2\right), \tag{1.2.9}$$

*where*

$$m^{2,\xi}_{t,s}(x) = x_2 + \int_t^s \Gamma_r x_1 dr + \int_t^s \bar{F}_2(\theta^2_{t,r}(\xi)) dr,$$

*and where $\bar{\Sigma}_{t,s}$ is the uniformly non-degenerate matrix given by (1.2.8).*

From this expression, we can give the following Gaussian type estimate on the transition density of the solution of the EDS (1.2.7) and on its derivatives :

**Lemma 1.2.2.** *Suppose that assumptions **(HL)** hold true, then, for $T$ small enough, the transition kernel $\bar{q}$ and its derivatives admit a Gaussian-type bound : there exists a positive constant $c$ depending only on known parameters in **(H)** such that for all $\xi$ in $\mathbb{R}^{2d}$ :*

$$\begin{aligned} &\left|D^{N^{x_1}}_{x_1} D^{N^{x_2}}_{x_2} D^{N^{y_1}}_{y_1} \bar{q}(t,x_1,x_2;s,y_1,y_2)\right| \qquad (1.2.10) \\ &\leq (s-t)^{-[3N^{x_2}+N^{x_1}+N^{y_1}]/2} \frac{c}{(s-t)^{2d}} \exp\left(-c\left(\frac{|y_1 - x_1|^2}{s-t} + \frac{|y_2 - m^{2,\xi}_{t,s}(x)|^2}{(s-t)^3}\right)\right), \end{aligned}$$

*for all $s$ in $(t,T]$, any $N^{x_1}, N^{x_2}, N^{y_1}$ less than 2.*

*Proof.* The case $N^{x_1} = N^{x_2} = N^{y_1} = 0$ in this his Lemma is shown in [DM10] (see Proposition 3.3), and we refer to Section 3 pp 18-24 of this paper for further details. Briefly, this result follows from the non-degeneracy of $\bar{\Sigma}_{0,1}$ and scaling properties of the system. We emphasize that the constant $c$ appearing in the exponential in estimate (1.2.10) does not depend on $\Gamma$, as suggested by Lemma 1.2.1. This uniform control is not obvious and is related to the "closed convex" assumption **(H3-b)**. We refer to Section 3 and Example

3.5 p14 of [DM10] for more details. The remainder of the proof is given in Section 1.3 in a more general setting, see Proposition 1.3.1. □

This Lemma says that each differentiation of the transition kernel w.r.t. the diffusive component gives a singularity of order $1/2$ while the differentiation w.r.t. the degenerate component gives a singularity of order $3/2$.

Note that for all $s$ in $[t,T]$, the mean $(x_1, m^{2,x}_{t,s}(x))$ of $\bar{X}^x_s$ satisfies the ODE (1.2.5) with initial condition $\xi = x$. Since this equation admits a unique solution under **(HL)**, we deduce that for all $s$ in $[t,T]$, the forward transport function is equal to the mean : $\theta_{t,s}(x) = m^x_{t,s}(x)$.

Finally, from the proof of Proposition 1.3.1 given in Section 1.3, one can also see that

$$\forall (t<s,x,y) \in [0,T]^2 \times \mathbb{R}^{2d} \times \mathbb{R}^{2d}, D_{x_2}\bar{q}(t,x;s,y) = -D_{y_2}\bar{q}(t,x;s,y). \tag{1.2.11}$$

This symmetry plays a crucial role in the proof of the Lipschitz estimates of the solution $u$ of the PDE (1.2.2) and of its derivative $D_{x_1}u$ .

### *1.2.3* Representation of the solution

The transition density (1.2.9) of the frozen process $\bar{X}$ is the fundamental solution of the heat equation :

$$\partial_t \bar{q}(t,x;T,y) + \bar{\mathcal{L}}\bar{q}(t,x;T,y) = 0,\ \bar{q}(T,x;T,y) = \delta_y(x),\ t \in [0,T],\ x,y \in \mathbb{R}^{2d},$$

where $\bar{\mathcal{L}} = (1/2)\Delta_{x_1} + \left[\bar{F}_2(\theta^2_{t,T}(\xi)) + \Gamma_T x_1\right] \cdot D_{x_2}$. Note that the PDE (1.2.2) reads :

$$\begin{aligned}&\partial_t u_i(t,x) + \bar{\mathcal{L}}u_i(t,x) = \phi_i(t,x) + (\bar{\mathcal{L}} - \mathcal{L})u_i(t,x),\ \text{for } (t,x) \in [0,T) \times \mathbb{R}^{2d}\\ &u_i(T,x) = 0, \quad i = 1,2.\end{aligned}$$

So that, for all $(t,x_1,x_2)$ in $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$, the unique solution of this PDE can be written as :

$$\begin{aligned}&u_i(t,x_1,x_2)\\ &= \int_t^T \int_{\mathbb{R}^{2d}} \phi_i(s,y_1,y_2)\bar{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds\\ &\quad - \int_t^T \int_{\mathbb{R}^{2d}} [\bar{F}_2(y_2) - \bar{F}_2(\theta^2_{t,T}(\xi))] \cdot D_{x_2}u_i(s,y_1,y_2)\bar{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds.\end{aligned}$$

### *1.2.4* A priori estimates

For sake of simplicity, we suppose throughout this subsection that $\bar{F}_2 \equiv 0$, so that

$$u_i(t,x_1,x_2) = \int_t^T \int_{\mathbb{R}^{2d}} \phi_i(s,y_1,y_2)\bar{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds. \tag{1.2.12}$$

In order to prove the Lipschitz estimates of $u$ and $D_{x_1}u$ of Proposition 1.1.3, we need to obtain estimates of the supremum norm of the first and second order derivatives of the $u_i$, $i=1,2$. As shown in Lemma 1.2.2, the differentiation of the transition density gives a time-singularity, so that it is not obvious that Lebesgue differentiation Theorem can be applied in (1.2.12).

Set $i$ in $\{1,2\}$ and let us focus on the worst case in Proposition 1.1.3, that is, the cross derivative $D_{x_1}D_{x_2}u_i$. Having in mind to use the Lebesgue differentiation Theorem, we focus on the cross differentiation of $\bar{q}$, which gives a time singularity of order 2. Let $(t,x_1,x_2)$ in

$[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$. In order to invert the integral and the differentiation operator, we split the integral : for any $\epsilon>0$, we can write :

$$\begin{aligned} u_i(t,x_1,x_2) &= \int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}}\phi_i(s,y_1,y_2)\bar q(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds\\ &\quad+\int_t^{t+\epsilon}\int_{\mathbb{R}^{2d}}\phi_i(s,y_1,y_2)\bar q(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds, \end{aligned}$$

and thanks to Lebesgue differentiation Theorem we have :

$$\begin{aligned} D_{x_1}D_{x_2}u_i(t,x_1,x_2) &= \int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}}\phi_i(s,y_1,y_2)D_{x_1}D_{x_2}\bar q(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds \quad (1.2.13)\\ &\quad+D_{x_1}D_{x_2}\left[\int_t^{t+\epsilon}\int_{\mathbb{R}^{2d}}\phi_i(s,y_1,y_2)\bar q(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds\right]. \end{aligned}$$

On a first hand, the last term in the right hand side reads [29] :

$$\begin{aligned} &D_{x_1}D_{x_2}\left[\int_t^{t+\epsilon}\int_{\mathbb{R}^{2d}}\phi_i(s,y_1,y_2)\bar q(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds\right]\\ &= D_{x_1}D_{x_2}\mathbb{E}\left[\int_t^{t+\epsilon}\phi_i(s,\bar X_s^{1,t,x},\bar X_s^{2,t,x})ds\right]. \end{aligned}$$

It follows from the proof of Lemma 1.1.2 in Section 1.1.6 that Lebesgue differentiation Theorem applies and there exists a positive constant $K$ such that

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^{2d}}\left|\mathbb{E}\left[\int_t^{t+\epsilon}D_{x_1}D_{x_2}\phi_i(s,\bar X_s^{1,t,x},\bar X_s^{2,t,x})ds\right]\right|\le K\epsilon. \quad (1.2.14)$$

On the other hand, for all $s$ in $(t,T]$ we have :

$$\begin{aligned} &\int_{\mathbb{R}^{2d}}\phi_i(s,y_1,y_2)D_{x_1}D_{x_2}\bar q(t,x_1,x_2;s,y_1,y_2)dy_1dy_2\\ &=\int_{\mathbb{R}^{2d}}\left(\phi_i(s,y_1,y_2)-\phi_i(s,y_1,\theta^2_{t,s}(\xi))\right)D_{x_1}D_{x_2}\bar q(t,x_1,x_2;s,y_1,y_2)dy_1dy_2\\ &\quad+\int_{\mathbb{R}^{2d}}\phi_i(s,y_1,\theta^2_{t,s}(\xi))D_{x_1}D_{x_2}\bar q(t,x_1,x_2;s,y_1,y_2)dy_1dy_2, \end{aligned}$$

Note that thanks to the symmetry (1.2.11), the last term in the right hand side is equal to 0. In the sequel, we refer to this argument as *the centering argument*. Combining this argument and the estimate for $D_{x_1}D_{x_2}\bar q$ in Lemma 1.2.2, we have, for $T$ small enough and all $s$ in $(t,T]$ :

$$\begin{aligned} &\int_{\mathbb{R}^{2d}}\left(\phi_i(s,y_1,y_2)-\phi_i(s,y_1,\theta^2_{t,s}(\xi))\right)D_{x_1}D_{x_2}\bar q(t,x_1,x_2;s,y_1,y_2)dy_1dy_2\\ &\le\int_{\mathbb{R}^{2d}}\left\{(s-t)^{-2}\left|\phi_i(s,y_1,y_2)-\phi_i(s,y_1,\theta^2_{t,s}(\xi))\right|\right.\\ &\qquad\left.\times\frac{c}{(s-t)^{2d}}\exp\left(-c\left(\frac{|y_1-x_1|^2}{s-t}+\frac{|y_2-m^{2,\xi}_{t,s}(x)|^2}{(s-t)^3}\right)\right)\right\}dy_1dy_2, \end{aligned}$$

29. The superscript "$t,x$" stands for the starting time and point of the process $\bar X$.

where $c$ depends only on known parameters in **(H)**. Using Hölder regularity of $\phi_i$ supposed in **(H1)**, we have :

$$\begin{aligned}
&\left|\int_{\mathbb{R}^{2d}}(s-t)^{-2}\left|\phi_i(s,y_1,y_2)-\phi_i(s,y_1,\theta^2_{t,s}(\xi))\right|\bar{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2\right|\\
&\quad\le C\int_{\mathbb{R}^{2d}}\Bigg\{(s-t)^{-2+3\beta^2_i/2}\frac{|y_2-\theta^2_{t,s}(\xi)|^{\beta^2_i}}{(s-t)^{3\beta^2_i/2}}\frac{c}{(s-t)^{2d}}\\
&\qquad\times\exp\left(-c\left(\frac{|y_1-x_1|^2}{s-t}+\frac{|y_2-m^{2,\xi}_{t,s}(x)|^2}{(s-t)^3}\right)\right)\Bigg\}dy_1dy_2.
\end{aligned}$$

Now, we use the off-diagonal decay of the Gaussian exponential : by letting $\xi=x$ (and then $\theta^2_{t,s}(x)=m^{2,x}_{t,s}(x)$), for all $\eta>0$, there exists a constant $\bar{C}>0$ such that [30]

$$\left(\frac{|y_2-m^{2,x}_{t,s}(x)|}{(s-t)^{3/2}}\right)^{\beta^2_i}\times\exp\left(-\eta\left(\frac{|y_1-x_1|^2}{s-t}+\frac{|y_2-m^{2,x}_{t,s}(x)|^2}{(s-t)^3}\right)\right)\le\bar{C},$$

where $\bar{C}$ depends on $\eta$ and $\beta^2_i$ only. Thus, by increasing the constant $c$ in the exponential, we obtain the following estimate :

$$\begin{aligned}
&\left|\int_{\mathbb{R}^{2d}}(s-t)^{-2}\left|\phi_i(s,y_1,y_2)-\phi_i(s,y_1,\theta^2_{t,s}(x))\right|\bar{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2\right| \qquad (1.2.15)\\
&\le C'\int_{\mathbb{R}^{2d}}(s-t)^{-2+3\beta^2_i/2}\frac{c}{(s-t)^{2d}}\exp\left(-c\left(\frac{|y_1-x_1|^2}{s-t}+\frac{|y_2-m^{2,\xi}_{t,s}(x)|^2}{(s-t)^3}\right)\right)dy_1dy_2.
\end{aligned}$$

Therefore, by choosing the value of $\beta^2_i>2/3$, the singularity $(s-t)^{-2+3\beta^2_i/2}$ becomes integrable. From (1.2.15), letting $\epsilon$ tends to 0 in (1.2.13) and using (1.2.14), we deduce that :

$$\|D_{x_1}D_{x_2}u_i\|_\infty\le C''T^{-1+3\beta^2_i/2}.$$

From this discussion, one can also see the specific choice of the freezing curve as the one that matches the off-diagonal decay of the exponential in $\bar{q}$ when $\xi=x$.

## *1.3* Mathematical tools

In this section, we introduce the ingredients for the proof of Proposition 1.1.3.

### *1.3.1* The frozen system

Consider the frozen system linearized around the transport of the initial condition (given as the solution of (1.3.2) below) :

$$\begin{aligned}
d\tilde{X}^{1,t,x}_s&=F_1(s,\theta_{t,s}(\xi))ds+\sigma(s,\theta_{t,s}(\xi))dW_s\\
d\tilde{X}^{2,t,x}_s&=F_2(s,\theta_{t,s}(\xi))ds+D_{x_1}F_2(s,\theta_{t,s}(\xi))(\tilde{X}^{1,t,x}_s-\theta^1_{t,s}(\xi))ds
\end{aligned}\qquad(1.3.1)$$

for all $s$ in $(t,T]$, any $t$ in $[0,T]$, and for any initial condition $x$ in $\mathbb{R}^{2d}$ at time $t$ and any $\xi\in\mathbb{R}^{2d}$ and where $(\theta_{t,s}(\xi))_{t\le s\le T}$ is defined by :

$$\frac{d}{ds}\theta_{t,s}(\xi)=F(s,\theta_{t,s}(\xi)),\quad\theta_{t,t}(\xi)=\xi.\qquad(1.3.2)$$

The following Proposition holds :

30. By using the inequality : $\forall\eta>0,\ \forall q>0,\ \exists\bar{C}>0$ s.t. $\forall\sigma>0,\ \sigma^qe^{-\eta\sigma}\le\bar{C}$.

**Proposition 1.3.1.** *Suppose that assumptions **(HR)** hold, then :*

*(i) There exists a unique (strong) solution of* (1.3.1) *with mean*

$$(m^{\xi}_{t,s})_{t\le s\le T} = (m^{1,\xi}_{t,s}, m^{2,\xi}_{t,s})_{t\le s\le T},$$

*where*

$$\begin{aligned} m^{1,\xi}_{t,s}(x) &= x_1 + \int_t^s F_1(r,\theta_{t,r}(\xi))dr, \qquad (1.3.3)\\ m^{2,\xi}_{t,s}(x) &= x_2 + \int_t^s \Big[F_2(r,\theta_{t,r}(\xi)) + D_{x_1}F_2(r,\theta_{t,r}(\xi))(x_1-\theta^1_{t,r}(\xi))\\ &\quad + D_{x_1}F_2(r,\theta_{t,r}(\xi))\int_t^r F_1(v,\theta_{t,v}(\xi))dv\Big]dr, \end{aligned}$$

*and uniformly non-degenerate covariance matrix* $(\tilde{\Sigma}_{t,s})_{t\le s\le T}$ *:*

$$\tilde{\Sigma}_{t,s} = \begin{pmatrix} \int_t^s \sigma\sigma^*(r,\theta_{t,r}(\xi))dr & \int_t^s R_{r,s}(\xi)\sigma\sigma^*(r,\theta_{t,r}(\xi))dr \\ \int_t^s \sigma\sigma^*(r,\theta_{t,r}(\xi))R^*_{r,s}(\xi)dr & \int_t^s R_{t,r}(\xi)\sigma\sigma^*(r,\theta_{t,r}(\xi))R^*_{t,r}(\xi)dr \end{pmatrix}, \qquad (1.3.4)$$

*where :*

$$R_{t,r}(\xi) = \left[\int_t^r D_{x_1}F_2(v,\theta_{t,v}(\xi))dv\right], \quad t\le r\le s\le T.$$

*(ii) This solution is a Gaussian process with transition density :*

$$\begin{aligned} &\tilde{q}(t,x_1,x_2;s,y_1,y_2) \qquad (1.3.5)\\ &= \frac{3^{d/2}}{(2\pi)^{d/2}}(\det[\tilde{\Sigma}_{t,s}])^{-1/2}\exp\left(-\frac{1}{2}|\tilde{\Sigma}^{-1/2}_{t,s}(y_1-m^{1,\xi}_{t,s}(x),y_2-m^{2,\xi}_{t,s}(x))^*|^2\right), \end{aligned}$$

*for all $s$ in $(t,T]$.*

*(iii) This transition density $\tilde{q}$ is the fundamental solution of the PDE driven by $\tilde{\mathcal{L}}^{t,\xi}$ and given by :*

$$\begin{aligned} \tilde{\mathcal{L}}^{t,\xi} &:= \frac{1}{2}Tr\left[a(t,\theta_{t,s}(\xi))D^2_{x_1}\right] + \left[F_1(t,\theta_{t,s}(\xi))\right]\cdot D_{x_1}\\ &\quad + \left[F_2(t,\theta_{t,s}(\xi)) + D_{x_1}F_2(t,\theta_{t,s}(\xi))\left(x_1-\theta^1_{t,s}(\xi)\right)\right]\cdot D_{x_2}. \qquad (1.3.6) \end{aligned}$$

*(iv) For $T$ small enough, there exist two positive constants $c$ and $C$, depending only on known parameters in **(H)**, such that*

$$\tilde{q}(t,x_1,x_2;s,y_1,y_2) \le C\hat{q}_c(t,x_1,x_2;s,y_1,y_2), \qquad (1.3.7)$$

*where*

$$\hat{q}_c(t,x_1,x_2;s,y_1,y_2) = \frac{c}{(s-t)^{2d}}\exp\left(-c\left(\frac{\left|y_1-m^{1,\xi}_{t,s}(x)\right|^2}{s-t} + \frac{\left|y_2-m^{2,\xi}_{t,s}(x)\right|^2}{(s-t)^3}\right)\right),$$

*and :*

$$\begin{aligned} &\left|D^{N^{x_1}}_{x_1}D^{N^{x_2}}_{x_2}D^{N^{y_1}}_{y_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right|\\ &\quad \le C(s-t)^{-[3N^{x_2}+N^{x_1}+N^{y_1}]/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2), \qquad (1.3.8) \end{aligned}$$

*for all $s$ in $(t,T]$ and any integers $N^{x_1}, N^{x_2}, N^{y_1}$ less than 2.*

*Proof.* (i) First of all, note that, under **(HR)**, the ODE : $[d/ds]\theta_{t,s}(\xi) = F(s,\theta_{t,s}(\xi))$, $\theta_{t,t}(\xi) = \xi$ admits a unique solution and (1.3.1) admits a unique strong solution $\tilde{X}$. One can write (1.3.1) as :

$$\begin{aligned}
\tilde{X}^{1,t,x}_s &= x_1 + \int_t^s F_1(r,\theta_{t,r}(\xi))dr + \int_t^s \sigma(r,\theta_{t,r}(\xi))dW_r,\\
\tilde{X}^{2,t,x}_s &= x_2 + \int_t^s \Big[F_2(r,\theta_{t,r}(\xi)) + D_{x_1}F_2(r,\theta_{t,r}(\xi))(x_1-\theta^1_{t,r}(\xi))\\
&\quad + D_{x_1}F_2(r,\theta_{t,r}(\xi))\int_t^r F_1(v,\theta_{t,v}(\xi))dv\Big]dr\\
&\quad + \int_t^s D_{x_1}F_2(r,\theta_{t,r}(\xi))\int_t^r \sigma(v,\theta_{t,v}(\xi))dW_v dr
\end{aligned}$$

Then, the expressions of the mean (1.3.3) and the variance (1.3.4) follow from the stochastic Fubini Theorem and standard computations. The uniform non-degeneracy of $(\tilde{\Sigma}_{t,s})_{t<s\leq T}$ arises from assumptions **(H)** and Proposition 3.1 in [DM10].

(ii)-(iii) These assertions result from standard computations.

(iv) For all $s$ in $(t,T]$, we know from Proposition 3.1 in [DM10] that the matrix $\tilde{\Sigma}_{t,s}$ is symmetric and uniformly non degenerate. Besides, from Subsection 2.3 and Proposition 3.4 in [DM10] there exists a constant $C$ depending only on known parameters in **(H)** such that : for all $s\in(t,T]$, for all $(x,y,\xi)\in\mathbb{R}^{2d}\times\mathbb{R}^{2d}\times\mathbb{R}^{2d}$,

$$\begin{aligned}
&-\left[\tilde{\Sigma}^{-1}_{t,s}(y_1-m^{1,\xi}_{t,s}(x), y_2-m^{2,\xi}_{t,s}(x))^*\right]\cdot\left[(y_1-m^{1,\xi}_{t,s}(x), y_2-m^{2,\xi}_{t,s}(x))^*\right]\\
&\quad\leq -C\left[\left(\frac{y_1-m^{1,\xi}_{t,s}(x)}{(s-t)^{1/2}}, \frac{y_2-m^{2,\xi}_{t,s}(x)}{(s-t)^{3/2}}\right)^*\right]\cdot\left[\left(\frac{y_1-m^{1,\xi}_{t,s}(x)}{(s-t)^{1/2}}, \frac{y_2-m^{2,\xi}_{t,s}(x)}{(s-t)^{3/2}}\right)^*\right].
\end{aligned}$$

For $i,j=1,2$, let $[\tilde{\Sigma}^{-1}_{t,s}]_{i,j}$ denotes the block of size $d\times d$ of the matrix $\tilde{\Sigma}^{-1}_{t,s}$ at the $(i-1)d+1,(j-1)d+1$ rank. We can deduce from (1.3.4) that there exists a positive constant $C$ depending only on known parameters in **(H)** such that (we also refer the reader to Lemma 3.6 and to the proof of Lemma 5.5 in [DM10] for more details), for all $s$ in $(t,T]$, for all $\zeta$ in $\mathbb{R}^d$ :

$$\begin{aligned}
\left|[\tilde{\Sigma}^{-1}_{t,s}]_{1,1}\zeta\right| &\leq C(s-t)^{-1}|\zeta|,\\
\left|[\tilde{\Sigma}^{-1}_{t,s}]_{1,2}\zeta\right| + \left|[\tilde{\Sigma}^{-1}_{t,s}]_{2,1}\zeta\right| &\leq C(s-t)^{-2}|\zeta|,\\
\left|[\tilde{\Sigma}^{-1}_{t,s}]_{2,2}\zeta\right| &\leq C(s-t)^{-3}|\zeta|,
\end{aligned}\tag{1.3.9}$$

hence, $\tilde{\Sigma}^{-1}_{t,\cdot}$ has the same structure as $K^{-1}_{\cdot-t}$ in (1.2.4).

Now, we compute the derivatives w.r.t. each component. Let $(t < s, x, y) \in [0,T]^2 \times \mathbb{R}^{2d} \times \mathbb{R}^{2d}$, we have :

$$
\begin{aligned}
&|D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)| \\
&= \left|\left(-2[\tilde{\Sigma}_{t,s}^{-1}]_{2,1}(y_1 - m_{t,s}^{1,\xi}(x)) - 2[\tilde{\Sigma}_{t,s}^{-1}]_{2,2}(y_2 - m_{t,s}^{2,\xi}(x))\right)\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right| \\
&\leq C(s-t)^{-3/2}\left(\left|\frac{(y_1 - m_{t,s}^{2,\xi}(x))}{(s-t)^{1/2}}\right| + \left|\frac{(y_2 - m_{t,s}^{2,\xi}(x))}{(s-t)^{3/2}}\right|\right)\tilde{q}(t,x_1,x_2;s,y_1,y_2) \\
&\leq C'(s-t)^{-3/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2).
\end{aligned}
$$

Note that the symmetry $D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2) = -D_{y_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)$ holds. Now, we have

$$
\begin{aligned}
&|D_{y_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)| \\
&= \left|\left(2[\tilde{\Sigma}_{t,s}^{-1}]_{1,1}(y_1 - m_{t,s}^{1,\xi}(x)) + 2[\tilde{\Sigma}_{t,s}^{-1}]_{1,2}(y_2 - m_{t,s}^{2,\xi}(x))\right)\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right| \\
&\leq C(s-t)^{-1/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2).
\end{aligned}
$$

Unfortunately, the transport of the initial condition of the diffusive component in the degenerate component breaks the symmetry and $D_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2) \neq -D_{y_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)$. Indeed

$$
\begin{aligned}
&D_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2) \\
&= \Bigg( -2[\tilde{\Sigma}_{t,s}^{-1}]_{1,1}(y_1 - m_{t,s}^{1,\xi}(x)) - 2[\tilde{\Sigma}_{t,s}^{-1}]_{1,2}(y_2 - m_{t,s}^{2,\xi}(x)) \\
&\qquad -2[\tilde{\Sigma}_{t,s}^{-1}]_{1,2}\left[(R_{t,s}(\xi))\,(y_1 - m_{t,s}^{1,\xi}(x))\right] \\
&\qquad -2[\tilde{\Sigma}_{t,s}^{-1}]_{2,2}\left[(R_{t,s}(\xi))\,(y_2 - m_{t,s}^{2,\xi}(x))\right]\Bigg)\tilde{q}(t,x_1,x_2;s,y_1,y_2).
\end{aligned}
$$

Since the term $R_{t,s}(\xi)$ is of order $(s-t)$ (this is the transport of the initial condition from time $t$ to $s$), we deduce that

$$
\begin{aligned}
&|D_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)| \\
&\leq C(s-t)^{-1/2}\left\{\left|\frac{(y_1 - m_{t,s}^{2,\xi}(x))}{(s-t)^{1/2}}\right| + \left|\frac{(y_2 - m_{t,s}^{2,\xi}(x))}{(s-t)^{3/2}}\right| + \left|\frac{(y_1 - m_{t,s}^{2,\xi}(x))}{(s-t)^{1/2}}\right|\right. \\
&\qquad \left. + \left|\frac{(y_2 - m_{t,s}^{2,\xi}(x))}{(s-t)^{3/2}}\right|\right\}\tilde{q}(t,x_1,x_2;s,y_1,y_2) \\
&\leq C'(s-t)^{-1/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
&D^2_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\\
&= \Big(-2[\tilde{\Sigma}^{-1}_{t,s}]_{1,1}D_{x_1}m^{1,\xi}_{t,s}(x)) - 2[\tilde{\Sigma}^{-1}_{t,s}]_{1,2}D_{x_1}m^{2,\xi}_{t,s}(x) - 2[\tilde{\Sigma}^{-1}_{t,s}]_{1,2}\left[(R_{t,s}(\xi))\,D_{x_1}m^{1,\xi}_{t,s}(x)\right]\\
&\quad -2[\tilde{\Sigma}^{-1}_{t,s}]_{2,2}\left[(R_{t,s}(\xi))\,D_{x_1}m^{2,\xi}_{t,s}(x)\right]\Big)\,\tilde{q}(t,x_1,x_2;s,y_1,y_2)\\
&\quad + \Big(-2[\tilde{\Sigma}^{-1}_{t,s}]_{1,1}(y_1 - m^{1,\xi}_{t,s}(x)) - 2[\tilde{\Sigma}^{-1}_{t,s}]_{1,2}(y_2 - m^{2,\xi}_{t,s}(x))\\
&\quad -2[\tilde{\Sigma}^{-1}_{t,s}]_{1,2}\left[(R_{t,s}(\xi))\,(y_1 - m^{1,\xi}_{t,s}(x))\right]\\
&\quad -2[\tilde{\Sigma}^{-1}_{t,s}]_{2,2}\left[(R_{t,s}(\xi))\,(y_2 - m^{2,\xi}_{t,s}(x))\right]\Big)^2\,\tilde{q}(t,x_1,x_2;s,y_1,y_2).
\end{aligned}
$$

Note that, from (1.3.3) we have $D_{x_1}m^{\xi}_{t,s}(x) = (\mathrm{Id}, R_{t,s}(\xi))^*$, so that,

$$
|D^2_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)| \le C(s-t)^{-1}\hat{q}_c(t,x_1,x_2;s,y_1,y_2).
$$

The other derivatives can be deduced from these computations and estimate (1.3.8) follows. $\square$

**Remark.** *From this proof, one can deduce that the symmetry $D_{x_2}\tilde{q} = -D_{y_2}\tilde{q}$ holds. Therefore, for all $t$ in $[0,T]$, all $s$ in $[t,T]$ and $y_1, x_1, x_2$ in $\mathbb{R}^d$,*

$$
\int_{\mathbb{R}^d} D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)dy_2 = 0. \tag{1.3.10}
$$

*This argument is very useful in the sequel.*

### *1.3.2* Definitions and rules of calculus

We introduce some definitions and rules of computations that will be usefull in the following section. Let us begin by the following definition :

**Definition 1.3.2.** *For all $\zeta$ in $\mathbb{R}^{2d}$ we denote by $\Delta(\zeta)$ the perturbation operator around $\zeta$ acting on any function $f$ from $[0,T]\times\mathbb{R}^{2d}$ as follows :*

$$
\forall (s,y) \in [0,T]\times\mathbb{R}^{2d},\ (\Delta(\zeta)f)(s,y) = f(s,y) - f(s,\zeta),
$$

*and for $j = 1,2$, for all $\zeta$ in $\mathbb{R}^{2d}$, we denote by $\Delta^j(\zeta)$ the perturbation operator around $\zeta_j$ acting on any function $f$ from $[0,T]\times\mathbb{R}^{2d}$ as follows :*

$$
\forall (s,y) \in [0,T]\times\mathbb{R}^{2d},\ (\Delta^1(\zeta)f)(s,y_1) = f(s,y_1,\zeta_2) - f(s,\zeta_1,\zeta_2),
$$

*and*

$$
\forall (s,y) \in [0,T]\times\mathbb{R}^{2d},\ (\Delta^2(\zeta)f)(s,y_1,y_2) = f(s,y_1,y_2) - f(s,y_1,\zeta_2).
$$

Given this definition we can give a generic *centering argument*, as introduce in the Brownian heuristic in Subsection 1.2 :

**Claim 1.3.3.** *Let $\tilde{q}$ be the function defined by (1.3.5) in Proposition 1.3.1 and let $f$ and $g$ be two continuous functions defined on $[0,T]\times\mathbb{R}^{2d}$.*

*Then, for all $N^1$ $N^2$ in $\mathbb{N}$, for all $t < s \in [0,T]^2$, $x\in\mathbb{R}^{2d}$ and $\zeta\in\mathbb{R}^{2d}$ we have that :*

*(a)* $D^{N^1}_{x_1}D^{N^2}_{x_2}\int_{\mathbb{R}^{2d}} f(s,y)\tilde{q}(t,x;s,y)dy = D^{N^1}_{x_1}D^{N^2}_{x_2}\int_{\mathbb{R}^{2d}}(\Delta(\zeta)f)(s,y)\tilde{q}(t,x;s,y)dy,$
*if $N_1 + N_2 > 0$,*

*(b)* $D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} f(s,y)\tilde{q}(t,x;s,y)dy = D_{x_1}^{N^1} D_{x_2}^{N^2} Q_{\tilde{q}} \int_{\mathbb{R}^{2d}} (\Delta^2(\zeta)f)(s,y)\tilde{q}(t,x;s,y)dy,$
*if* $N_2 > 0$,

*(c)* $D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} (\Delta^1(\zeta)f)(s,y)g(s,y)\tilde{q}(t,x;s,y)dy$
$= D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} (\Delta^1(\zeta)f)(s,y)(\Delta^2(\zeta)g)(s,y)\tilde{q}(t,x;s,y)dy,$
*if* $N_2 > 0$.

*Proof.* Let $f$ and $g$ be defined as in Claim 1.3.3 and let $t < s \in [0,T]^2$, $x \in \mathbb{R}^{2d}$. We have, by Definition 1.3.2 of $\Delta$ :

$$\begin{aligned} D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} f(s,y)\tilde{q}(t,x;s,y)dy &= D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} (\Delta(\zeta)f)(s,y)\tilde{q}(t,x;s,y)dy \\ &+ D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} f(s,\zeta)\tilde{q}(t,x;s,y)dy, \end{aligned}$$

for all $\zeta$ in $\mathbb{R}^{2d}$. The last term in the right hand side is equal to 0 since it does not depend on $x$. This concludes the proof of (a). Now, we prove (c). For all $\zeta$ in $\mathbb{R}^{2d}$ we have

$$\begin{aligned} &D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} (\Delta^1(\zeta)f)(s,y_1)g(s,y_1,y_2)\tilde{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2 \\ &= D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} (\Delta^1(\zeta)f)(s,y_1)(\Delta^2(\zeta)g)(s,y_1,y_2)\tilde{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2 \\ &\quad + D_{x_1}^{N^1} D_{x_2}^{N^2} \int_{\mathbb{R}^{2d}} (\Delta^1(\zeta)f)(s,y_1)g(s,y_1,\zeta_2)\tilde{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2. \end{aligned}$$

By using again Lebesgue differentiation Theorem, Remark 1.3.10 (since $N^2 > 0$) and integration by parts w.r.t. "$y_2$", the last term in the right hand side is equal to 0. Finally, assertion (b) follows from the same arguments. This concludes the proof of the Claim. ☐

### *1.3.3* Representation and differentiation of the solution of the PDE (1.1.9)

**Lemma 1.3.4.** *Suppose that assumptions* ***(HR)*** *hold, then, the solution* $u = (u_1, u_2)^*$ *of the PDE* (1.1.9) *can be written as, for all* $x \in \mathbb{R}^{2d}$ *and* $t \in [0,T]$ *:*

$$u_i(t,x) = \int_t^T \mathbb{E}\left[\phi_i(s,\tilde{X}_s^{t,x}) - (\mathcal{L} - \tilde{\mathcal{L}}^{t,\xi})u_i(s,\tilde{X}_s^{t,x})\right] ds, \tag{1.3.11}$$

*for all* $\xi \in \mathbb{R}^{2d}$*, for* $i = 1,2$*, where* $X^{t,x} = (X^{2,t,x}, X^{1,t,x})^*$ *is the solution of the frozen system* (1.3.1) *with initial condition* $x$ *at time* $t$*, the operator* $\tilde{\mathcal{L}}^{t,\xi}$ *is given in* (1.3.6)*. Note* *that :*

$$\begin{aligned}
u_i(t,x) &= \int_t^T \int_{\mathbb{R}^{2d}} \phi_i(s,y)\tilde{q}(t,x;s,y)dyds \\
&\quad - \int_t^T \int_{\mathbb{R}^{2d}} \frac{1}{2}\mathrm{Tr}\left[\Delta(\theta_{t,s}(\xi))a(s,y)D^2_{x_1}u_i(s,y)\right]\tilde{q}(t,x;s,y)dyds \\
&\quad - \int_t^T \int_{\mathbb{R}^{2d}} [\Delta(\theta_{t,s}(\xi))F_1(s,y)]\cdot D_{x_1}u_i(s,y)\tilde{q}(t,x;s,y)dyds \\
&\quad - \int_t^T \int_{\mathbb{R}^{2d}} \left[\Delta(\theta_{t,s}(\xi))F_2(s,y) - D_{x_1}F_2(s,\theta_{t,s}(\xi))\Delta^1(\theta_{t,s}(\xi))y_1\right] \\
&\qquad \cdot D_{x_2}u_i(s,y)\tilde{q}(t,x;s,y)dyds \\
&:= H_i^1(t;t,x) + H_i^2(t;t,x) + H_i^3(t;t,x) + H_i^4(t;t,x), \qquad (1.3.12)
\end{aligned}$$

*Moreover, this solution is infinitely differentiable and for all* $\xi \in \mathbb{R}^{2d}$*, for all* $\epsilon > 0$*, all* $N^1$*,* $N^2$ *in* $\mathbb{N}^*$*, we have :*

$$D^{N^1}_{x_1}D^{N^2}_{x_2}u_i(t,x) = \int_{t+\epsilon}^T \mathbb{E}\left[\phi_i(s,\tilde{X}_s^{t,x}) - (\mathcal{L}-\tilde{\mathcal{L}}^{t,\xi})u_i(s,\tilde{X}_s^{t,x})\right]ds + \mathcal{O}(\epsilon) \qquad (1.3.13)$$

*for* $i = 1, 2$*. We also write*

$$u_i(t,x) = H_i^1(t+\epsilon;t,x) + H_i^2(t+\epsilon;t,x) + H_i^3(t+\epsilon;t,x) + H_i^4(t+\epsilon;t,x) + \mathcal{O}(\epsilon),$$

*where the terms* $H_i^j(t+\epsilon;t,x)$*,* $j = 1, \cdots, 4$ *are properly defined by identifying* (1.3.13) *with* (1.3.12)*.*

*Proof.* We recall that from Lemma 1.1.2, the PDE (1.1.9) is well posed and can be rewritten as :

$$\begin{aligned}
&\partial_t u_i(t,x) + \tilde{\mathcal{L}}^{t,\xi}u_i(t,x) = -(\mathcal{L}u_i(t,x) - \tilde{\mathcal{L}}^{t,\xi}u_i(t,x)) + \phi_i(t,x),\ (t,x) \in [0,T) \times \mathbb{R}^{2d} \\
&u_i(T,x) = 0,\ i = 1, 2. \qquad (1.3.14)
\end{aligned}$$

From Feynman-Kac representation, the solution of (1.3.14) reads :

$$u_i(t,x) = \int_t^T \mathbb{E}\left[\phi_i(s,\tilde{X}_s^{t,x}) - (\mathcal{L}-\tilde{\mathcal{L}}^{t,\xi})u_i(t,\tilde{X}_s^{t,x})\right]ds,$$

for all $(t,x) \in [0,T] \times \mathbb{R}^{2d}$. This gives the first term in the right hand side of (1.3.12). Next, given $\epsilon > 0$, we have

$$\begin{aligned}
u_i(t,x) &= \int_{t+\epsilon}^T \mathbb{E}\left[\phi_i(s,\tilde{X}_s^{t,x}) - (\mathcal{L}-\tilde{\mathcal{L}}^{t,\xi})u_i(s,\tilde{X}_s^{t,x})\right]ds \\
&\quad + \int_t^{t+\epsilon} \mathbb{E}\left[\phi_i(s,\tilde{X}_s^{t,x}) - (\mathcal{L}-\tilde{\mathcal{L}}^{t,\xi})u_i(s,\tilde{X}_s^{t,x})\right]ds.
\end{aligned}$$

Under **(HR)**, the coefficients of $\mathcal{L}$, $\tilde{\mathcal{L}}^{t,\xi}$ and the functions $\phi_i$, $u_i$ are smooth. So that, from the proof of Lemma 1.1.2 in Subsection 1.1.6, we deduce that there exists a positive constant $K$, depending on the mollifying procedure, such that, for all $s$ in $(t,T]$ :

$$\left|D^{N^1}_{x_1}D^{N^2}_{x_2}\mathbb{E}\left[\phi_i(s,\tilde{X}_s^{t,x}) - (\mathcal{L}-\tilde{\mathcal{L}}^{t,\xi})u_i(s,\tilde{X}_s^{t,x})\right]\right| \leq K.$$

The claim follows from Lebesgue differentiation Theorem. $\square$

## *1.4* Proof of Proposition 1.1.3

### *1.4.1* From parametrix to uniform Lipschitz estimates

We give preliminary results in order to prove the Lipschitz bounds. These bounds are obtained under **(HR)** but depend only on known parameters in **(H)**.

**Proof of Proposition 1.1.3.** We prove Proposition 1.1.3 by using a circular argument since the representation (1.3.12) of each $u_i$, $i = 1, 2$ involves the derivatives themselves. In the following, $u_i$ denotes the $i^{th}$ component of the solution $u = (u_1, u_2)^*$ of the linear system of PDE (1.1.9). The following Lemmas hold for $i = 1, 2$.

**Lemma 1.4.1.** *Suppose assumptions **(HR)** hold. Then, for $T$ small enough there exist two positive numbers $\delta_{1.4.1}$ and $\bar{\delta}_{1.4.1}$ and a positive constant $C$ depending only on known parameters in **(H)** such that :*

$$\left\|D_{x_1}^2 u_i\right\|_\infty \le T^{\delta_{1.4.1}} C \left(1 + \|D_{x_2} u_i\|_\infty\right),$$

*and*

$$\|D_{x_1} u_i\|_\infty \le T^{\bar{\delta}_{1.4.1}} C \left(1 + \|D_{x_2} u_i\|_\infty\right).$$

**Lemma 1.4.2.** *Suppose assumptions **(HR)** hold. Then, for $T$ small enough, there exist a $\delta_{1.4.2} > 0$ and a positive constant $C$ depending only on known parameters in **(H)**, such that :*

$$\|D_{x_2} u_i\|_\infty \le C T^{\delta_{1.4.2}} \left(1 + \|D_{x_1} D_{x_2} u_i\|_\infty\right).$$

**Lemma 1.4.3.** *Suppose assumptions **(HR)** hold. Then, for $T$ small enough, there exist a positive real $\delta_{1.4.3}$ and a positive constant $C$, depending only on known parameters in **(H)** such that :*

$$\|D_{x_1} D_{x_2} u_i\|_\infty \le C T^{\bar{\delta}_{1.4.3}}.$$

Proposition 1.1.3 follows from Lemmas 1.4.1, 1.4.2 and 1.4.3. □

### *1.4.2* Proof of Lemma 1.4.1

Here we prove that for $T$ small enough, there exists a positive constant $C$ depending only on known parameters in **(H)** such that :

$$\text{(i)} \quad \left\|D_{x_1}^2 u_i\right\|_\infty \le C \Big\{ \left(T^{3\beta_2^2/2} + T^{(1+\alpha^1)/2}\right) \|D_{x_2} u_i\|_\infty + \left(T^{\beta_1^1/2} + T^{3\beta_1^2/2}\right) \|D_{x_1} u_i\|_\infty + T^{\beta_i^1/2} + T^{3\beta_i^2/2} \Big\},$$

$$\text{(ii)} \quad \|D_{x_1} u_i\|_\infty \le C \Big\{ \left(T^{1/2+3\beta_2^2/2} + T^{(3+\alpha^1)/2}\right) \|D_{x_2} u_i\|_\infty + \left(T + T^2\right) \left\|D_{x_1}^2 u_i\right\|_\infty + T^{(1+\beta_i^1)/2} + T^{(1+3\beta_i^2)/2} \Big\}.$$

We first show (i). Let $\epsilon > 0$, from the representation (1.3.12) and Lemma 1.3.4 we have :

$$D_{x_1}^2 u_i(t, x_1, x_2) = \sum_{j=1}^{4} D_{x_1}^2 H_i^j(t + \epsilon; t, x_1, x_2) + \mathcal{O}(\epsilon), \tag{1.4.1}$$

and thanks to assertion (a) in Claim 1.3.3 :

$$
\begin{aligned}
&D^2_{x_1} u_i(t,x_1,x_2)\\
&= \int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \Delta(\theta_{t,s}(\xi))\phi_i(s,y_1,y_2)\left[D^2_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right] dy_1dy_2ds\\
&\quad -\frac{1}{2}\int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \mathrm{Tr}\left[\Delta(\theta_{t,s}(\xi))a(s,y_1,y_2)D^2_{x_1}u_i(s,y_1,y_2)\right]\left[D^2_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right] dy_1dy_2ds\\
&\quad -\int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \left[\Delta(\theta_{t,s}(\xi))F_1(s,y_1,y_2)\right]\cdot D^2_{x_1}u_i(s,y_1,y_2)\left[D^2_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right] dy_1dy_2ds\\
&\quad +\int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \left[\Delta^2(\theta_{t,s}(\xi))F_2(s,y_1,y_2) - D_{x_1}F_2(s,\theta^1_{t,s}(\xi),\theta^2_{t,s}(\xi))\Delta(\theta_{t,s}(\xi))y_1\right]\\
&\quad \cdot\left[D_{x_2}u_i(s,y_1,y_2)\right]\left[D^2_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right] dy_1dy_2ds + \mathcal{O}(\epsilon).
\end{aligned}
$$

Then, note that thanks to a Taylor expansion of order 0 with integrable remainder of the mapping $y_1\in\mathbb{R}^d \mapsto F_2(s,y_1,\theta^2_{t,s}(\xi))$ around $\theta^1_{t,s}(\xi)$ and **(H3-a)**, we have that :

$$
\begin{aligned}
&\left|\left[\Delta(\theta_{t,s}(\xi))F_2(s,y) - D_{x_1}F_2(s,\theta_{t,s}(\xi))\left(\Delta^1(\theta_{t,s}(\xi))y_1\right)\right]\cdot D_{x_2}u_i(s,y)\right|\\
&\quad \le C\left\|D_{x_2}u_i\right\|_\infty\left(|\Delta^2(\theta_{t,s}(\xi))y_2|^{\beta^2_2} + |\Delta^1(\theta_{t,s}(\xi))y_1|^{1+\alpha^1}\right). \qquad (1.4.2)
\end{aligned}
$$

From Proposition 1.3.1, we know that for all $s$ in $(t,T]$ and $y\in\mathbb{R}^{2d}$, $D^2_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\le C'(s-t)^{-1}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)$. Combining this argument with the regularity of the coefficients given in **(H)**, and thanks to (1.4.2), we obtain :

$$
\begin{aligned}
&\left|D^2_{x_1}u_i(t,x_1,x_2)\right|\\
&\le C''\int_{t+\epsilon}^T (s-t)^{-1}\int_{\mathbb{R}^{2d}}\left\{\sum_{j=1}^2\left\{(s-t)^{(j-1/2)\beta^j_i}\left|\frac{\Delta^j(\theta_{t,s}(\xi))y_j}{(s-t)^{(j-1/2)}}\right|^{\beta^j_i}\right.\right.\\
&\quad \left. +\left\|D^2_{x_1}u_i\right\|_\infty (s-t)^{(j-1/2)}\left|\frac{\Delta^j(\theta_{t,s}(\xi))y_j}{(s-t)^{(j-1/2)}}\right| + \left\|D_{x_1}u_i\right\|_\infty (s-t)^{(j-1/2)\beta^j_1}\left|\frac{\Delta^j(\theta_{t,s}(\xi))y_j}{(s-t)^{(j-1/2)}}\right|^{\beta^j_1}\right\}\\
&\quad \left. +\left\|D_{x_2}u_i\right\|_\infty\left[(s-t)^{3\beta^2_2/2}\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|^{\beta^2_2} + (s-t)^{(1+\alpha^1)/2}\left|\frac{\Delta^1(\theta_{t,s}(\xi))y_1}{(s-t)^{1/2}}\right|^{1+\alpha^1}\right]\right\}\\
&\quad \times\hat{q}_c(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds + \mathcal{O}(\epsilon).
\end{aligned}
$$

Set $\xi = x$, by using the off-diagonal decay of the Gaussian exponential in $\hat{q}_c$ (see the computations in Subsection 1.2.4) and by integrating w.r.t the space variables we have :

$$
\begin{aligned}
&\left|D^2_{x_1}u_i(t,x_1,x_2)\right|\\
&\le C'''\int_{t+\epsilon}^T\left\{\sum_{j=1}^2\left\{(s-t)^{-1+(j-1/2)\beta^j_i} + \left((s-t)^{-1/2}+(s-t)^{1/2}\right)\left\|D^2_{x_1}u_i\right\|_\infty\right.\right.\\
&\quad \left.\left. +(s-t)^{(j-1/2)\beta^j_2-1}\left\|D_{x_1}u_i\right\|_\infty\right\} + \left((s-t)^{-1+3\beta^2_1/2}+(s-t)^{(\alpha^1-1)/2}\right)\left\|D_{x_2}u_i\right\|_\infty\right\}ds\\
&\quad +\mathcal{O}(\epsilon).
\end{aligned}
$$

By letting $\epsilon\to 0$, we finally obtain :

$$\left\|D_{x_1}^2 u_i\right\|_\infty \leq C'''\Big\{ \left(T^{3\beta_2^2/2} + T^{(1+\alpha^1)/2}\right) \|D_{x_2} u_i\|_\infty + \left(T^{\beta_1^1/2} + T^{3\beta_1^2/2}\right) \|D_{x_1} u_i\|_\infty$$
$$+ \left(T^{1/2} + T^{3/2}\right) \left\|D_{x_1}^2 u_i\right\|_\infty + T^{\beta_i^1/2} + T^{3\beta_i^2/2}\Big\}.$$

Then, taking $T$ small enough such that $C'''(T^{1/2}+T^{3/2}) = 1/2$ we deduce the assertion (i) from a circular argument. The proof of the second statement (ii) can be done by the same arguments. This concludes the proof of Lemma 1.4.1. □

### *1.4.3* Proof of Lemma 1.4.2

We first derive a representation formula for $H_i^2$ and $H_i^3$ defined by (1.3.12) to handle the singularity of the derivative of the kernel $\tilde{q}$. These formulas are given in the following claim.

**Claim 1.4.4.** *For all $(t, x_1, x_2) \in [0, T] \times \mathbb{R}^{2d}$, for all $\epsilon > 0$, we have :*

$$\begin{aligned}
&D_{x_2} H_i^2(t+\epsilon; t, x_1, x_2) \qquad (1.4.3)\\
&= -\frac{1}{2}\int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \mathrm{Tr}\left[[\Delta^2(\theta_{t,s}(\xi)) a(s, y_1, y_2)] D_{x_1}^2 u_i(s, y_1, y_2)\right] D_{x_2}\tilde{q}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds\\
&\quad + \frac{1}{2}\int_{t+\epsilon}^T \Big\{ \Big[ \sum_{l=1}^d \int_{\mathbb{R}^{2d}} \left[\frac{\partial}{\partial y_{1l}} a_{l.}(s, y_1, \theta_{t,s}^2(\xi))\right] \cdot [\Delta^2(\theta_{t,s}(\xi)) D_{x_1} u_i(s, y_1, y_2)] \Big]\\
&\qquad \times [D_{x_2}\tilde{q}(t, x_1, x_2; s, y_1, y_2)] \Big\} dy_1 dy_2 ds\\
&\quad + \frac{1}{2}\sum_{l=1}^d \int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} [\Delta^1(\theta_{t,s}(\xi)) a_{l.}(s, y_1, \theta_{t,s}^2(\xi))] \cdot [\Delta^2(\theta_{t,s}(\xi)) D_{x_1} u_i(s, y_1, y_2)]\\
&\qquad \times \left[ D_{x_2}\left(\frac{\partial}{\partial y_{1l}} \tilde{q}(t, x_1, x_2; s, y_1, y_2)\right)\right] dy_1 dy_2 ds,
\end{aligned}$$

*where "$a_{l.}$" denotes the $l^{\text{th}}$ line of the matrix $a$, and*

$$\begin{aligned}
&D_{x_2} H_i^3(t+\epsilon; t, x_1, x_2) \qquad (1.4.4)\\
&= -\int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \Big\{ \Big( [\Delta^2(\theta_{t,s}(\xi)) F_1(s, y_1, y_2)] \cdot D_{x_1} u_i(s, y_1, y_2) \Big)\\
&\qquad \times D_{x_2}\tilde{q}(t, x_1, x_2; s, y_1, y_2) \Big\} dy_1 dy_2 ds\\
&\quad - \int_{t+\epsilon}^T \Big\{ \int_{\mathbb{R}^{2d}} \Big( [\Delta^1(\theta_{t,s}(\xi)) F_1(s, y_1, \theta_{t,s}^2(\xi))] \cdot [\Delta^2(\theta_{t,s}(\xi)) D_{x_1} u_i(s, y_1, y_2)] \Big)\\
&\qquad \times D_{x_2}\tilde{q}(t, x_1, x_2; s, y_1, y_2) \Big\} dy_1 dy_2 ds.
\end{aligned}$$

*Proof of Claim 1.4.4 :* Start with (1.4.3). Since by Definition 1.3.2 of $\Delta$ we have $\Delta = \Delta^1 + \Delta^2$, by an integration by parts we get

$$\begin{aligned}
&H_i^2(t+\epsilon;t,x_1,x_2)\\
&= -\int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \left( \frac{1}{2}\mathrm{Tr}\left[ [\Delta^2(\theta_{t,s}(\xi)) a(s,y_1,y_2)] D^2_{y_1} u_i(s,y_1,y_2) \right] \right) \tilde{q}(t,x_1,x_2;s,y_1,y_2) dy_1 dy_2 ds\\
&\quad + \frac{1}{2}\sum_{l=1}^d \int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \left[ \frac{\partial}{\partial y_{1l}} \Delta^1(\theta_{t,s}(\xi)) a_{l.}(s,y_1,\theta^2_{t,s}(\xi)) \right] \cdot D_{y_1} u_i(s,y_1,y_2) \tilde{q}(t,x_1,x_2;s,y_1,y_2) dy_1 dy_2 ds\\
&\quad + \frac{1}{2}\sum_{l=1}^d \int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \left[ \Delta^1(\theta_{t,s}(\xi)) a_{l.}(s,y_1,\theta^2_{t,s}(\xi)) \right] \cdot D_{y_1} u_i(s,y_1,y_2) \frac{\partial}{\partial y_{1l}} \tilde{q}(t,x_1,x_2;s,y_1,y_2) dy_1 dy_2 ds.
\end{aligned}$$

Note that, for all $l \in \{1,\cdots,d\}$, $[\partial/\partial y_{1l}]\Delta^1(\theta_{t,s}(\xi)) a_{l.}(s,y_1,\theta^2_{t,s}(\xi)) = [\partial/\partial y_{1l}] a_{l.}(s,y_1,\theta^2_{t,s}(\xi))$. We conclude by applying assertion (c) of Claim 1.3.3 to the two last terms in the right hand side with $D_{x_1} u_i$ as $g$.

Assertion (1.4.4) follows from of Claim 1.3.3. This concludes the proof of Claim 1.4.4. □

It follows from the representation (1.3.12) and Lemma 1.3.4 that, for all $\epsilon > 0$ :

$$D_{x_2} u_i(t,x_1,x_2) = \sum_{j=1}^4 D_{x_2} H_i^j(t+\epsilon;t,x_1,x_2) + \mathcal{O}(\epsilon) \tag{1.4.5}$$

We bound each $D_{x_2} H_i^j$, $j = 1\ldots,4$. We recall from Proposition 1.3.1 that there exists a positive constant $C$ depending only on known parameters in **(H)** such that, for all $s$ in $(t,T]$ and $(y_1,y_2) \in \mathbb{R}^{2d}$ :

$$D_{x_2} \tilde{q}(t,x_1,x_2;s,y_1,y_2) \leq C(s-t)^{-3/2} \hat{q}_c(t,x_1,x_2;s,y_1,y_2). \tag{1.4.6}$$

*Bound of $D_{x_2} H_i^1$.* By applying assertion (b) of Claim 1.3.3 and thanks to **(H)** and (1.4.6) :

$$\begin{aligned}
&\left| D_{x_2} H_i^1(t+\epsilon;t,x_1,x_2) \right|\\
&= \left| \int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \Delta^2(\theta_{t,s}(\xi)) \phi_i(s,y_1,y_2) \left[ D_{x_2} \tilde{q}(t,x_1,x_2;s,y_1,y_2) \right] dy_1 dy_2 ds \right|\\
&\leq C' \int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} (s-t)^{-3/2(1-\beta_i^2)} \left| \frac{\Delta^2(\theta_{t,s}(\xi)) y_2}{(s-t)^{3/2}} \right|^{\beta_i^2} \hat{q}_c(t,x_1,x_2;s,y_1,y_2) dy_1 dy_2 ds\\
&\leq C'' \int_{t+\epsilon}^T (s-t)^{-3/2(1-\beta_i^2)} ds.
\end{aligned} \tag{1.4.7}$$

*Bound of $D_{x_2} H_i^2$.* By using Mean Value Theorem, we have, for all $(s,y_1,y_2) \in (t,T] \times \mathbb{R}^d \times \mathbb{R}^d$ :

$$\left| \Delta^2(\theta_{t,s}(\xi)) D_{x_1} u_i(s,y_1,y_2) \right| \leq \| D_{x_1} D_{x_2} u_i \|_\infty \left| \Delta^2(\theta_{t,s}(\xi)) y_2 \right|, \tag{1.4.8}$$

and from Proposition 1.3.1 we have : for all $l$ in $\{1,\cdots,d\}$

$$\left| D_{x_2}\left( [\partial/\partial y_{1l}] \tilde{q}(t,x_1,x_2;s,y_1,y_2) \right) \right| \leq C'(s-t)^{-2} \hat{q}_c(t,x_1,x_2;s,y_1,y_2).$$

Hence, from representation (1.4.3) of $H_i^2$, Lipschitz regularity of $a$ from **(H)**, estimates (1.4.6) and (1.4.8) we have that :

$$\begin{aligned}
&\left|D_{x_2}H_i^2(t+\epsilon;t,x_1,x_2)\right| \\
&\le C''\int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}}\Bigg\{\left\|D_{x_1}^2u_i\right\|_\infty\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|+\|D_{x_1}D_{x_2}u_i\|_\infty\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right| \\
&\quad+\|D_{x_1}D_{x_2}u_i\|_\infty\left|\frac{\Delta^1(\theta_{t,s}(\xi))y_1}{(s-t)^{1/2}}\right|\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|\Bigg\}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds \\
&\le C'''\int_{t+\epsilon}^T\left(\left\|D_{x_1}^2u_i\right\|_\infty+\|D_{x_1}D_{x_2}u_i\|_\infty\right)ds,
\end{aligned}\tag{1.4.9}$$

by letting $\xi=x$ and by using the diagonal decay of $\hat{q}_c$.

*Bound of $D_{x_2}H_i^3$.* From representation (1.4.4) of $D_{x_2}H_i^3$ and then by using the regularity of $F_1$, and estimates (1.4.6) and (1.4.8),

$$\begin{aligned}
&\left|D_{x_2}H_i^3(t+\epsilon;t,x_1,x_2)\right| \\
&\le C'\int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}}\Bigg\{(s-t)^{-3/2(1-\beta_1^2)}\|D_{x_1}u_i\|_\infty\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|^{\beta_1^2}+C'\left\|D_{x_1}D_{x_2}u_i\right\|_\infty \\
&\quad\times(s-t)^{\beta_1^1/2}\left|\frac{\Delta^1(\theta_{t,s}(\xi))y_1}{(s-t)^{1/2}}\right|^{\beta_1^1}\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|\Bigg\}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds.
\end{aligned}\tag{1.4.10}$$

By setting $\xi=x$, we obtain

$$\begin{aligned}
&\left|D_{x_2}H_i^3(t+\epsilon;t,x_1,x_2)\right| \\
&\le C''\int_{t+\epsilon}^T\|D_{x_1}u_i\|_\infty(s-t)^{-3/2(1-\beta_1^2)}+\|D_{x_1}D_{x_2}u_i\|_\infty(s-t)^{\beta_1^1/2}ds.
\end{aligned}\tag{1.4.11}$$

*Bound of $D_{x_2}H_i^4$.* From (1.3.12) and by using the regularity of $F_2$, (1.4.2) and (1.4.6), we have :

$$\begin{aligned}
&\left|D_{x_2}H_i^4(t+\epsilon;t,x_1,x_2)\right| \\
&=\Bigg|\int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}}\Bigg\{\left[\Delta^1(\theta_{t,s}(\xi))F_2(s,y_1,\theta_{t,s}^2(\xi))-D_{x_1}F_2(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi))\Delta^1(\theta_{t,s}(\xi))y_1\right] \\
&\quad\cdot[D_{x_2}u_i(s,y_1,y_2)]-\left[\Delta^2(\theta_{t,s}(\xi))F_2(s,y_1,y_2)\right]\cdot D_{x_2}u_i(s,y_1,y_2)\Bigg\}D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds\Bigg| \\
&\le C'\left\|D_{x_2}u_i\right\|_\infty\int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}}\Bigg\{(s-t)^{-1+\alpha^1/2}\left|\frac{\Delta^1(\theta_{t,s}(\xi))y_1}{(s-t)^{1/2}}\right|^{1+\alpha^1} \\
&\quad+(s-t)^{-3(1-\beta_2^2)/2}\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|^{\beta_2^2}\Bigg\}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds.
\end{aligned}$$

By setting $\xi = x$ we obtain

$$\begin{aligned}&\left|D_{x_2} H_i^4(t, t+\epsilon, x_1, x_2)\right| \\ &\leq C'' \int_{t+\epsilon}^{T} \|D_{x_2} u_i\|_\infty \left((s-t)^{-1+\alpha^1/2} + (s-t)^{-3(1-\beta_2^2)/2}\right) ds. \end{aligned} \tag{1.4.12}$$

By letting $\epsilon$ tends to 0 in (1.4.7), (1.4.9), (1.4.11), (1.4.12), by combining these estimates and using estimates of $\|D_{x_1} u_i\|_\infty$ and $\|D_{x_1}^2 u_i\|_\infty$ from Lemma 1.4.1, we can find a positive number $\delta_{1.4.2}$ such that, for $T$ small enough :

$$\|D_{x_2} u_i\|_\infty \leq CT^{\delta_{1.4.2}} \left(1 + \|D_{x_1} D_{x_2} u_i\|_\infty\right),$$

where $C$ and $\delta_{1.4.2}$ depend only on known parameters in **(H)**. This concludes the proof of Lemma 1.4.2. □

### *1.4.4* Proof of Lemma 1.4.3

We first derive a representation formula for the term $H_i^4$ in (1.3.12) to handle the singularity of the derivative of the kernel $\tilde{q}$ : as a consequence of Claim 1.3.3 and definition of $H_i^4$ we have

**Claim 1.4.5.** *For all $\epsilon > 0$, we have :*

$$\begin{aligned}&D_{x_2} H_i^4(t+\epsilon; t, x_1, x_2) = \\ &- \int_{t+\epsilon}^{T} \int_{\mathbb{R}^{2d}} \Big\{ \left[\Delta^1(\theta_{t,s}(\xi)) F_2(s, y_1, \theta_{t,s}^2(\xi)) - D_{x_1} F_2(s, \theta_{t,s}^1(\xi), \theta_{t,s}^2(\xi)) \Delta^1(\theta_{t,s}(\xi)) y_1\right] \\ &\quad \cdot \left[\Delta^2(\theta_{t,s}(\xi)) D_{x_2} u_i(s, y_1, y_2)\right] \Big\} D_{x_2} \tilde{q}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds \\ &- \int_{t+\epsilon}^{T} \int_{\mathbb{R}^{2d}} \left[\Delta^2(\theta_{t,s}(\xi)) F_2(s, y_1, y_2)\right] \cdot D_{x_2} u_i(s, y_1, y_2) D_{x_2} \tilde{q}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds,\end{aligned} \tag{1.4.13}$$

In order to prove Lemma 1.4.3, we need to obtain an estimate on the regularity of $D_{x_2} u_i$ w.r.t. $x_2$. This estimate is given in the following claim.

**Claim 1.4.6.** *Under the assumptions of Lemma 1.4.2, for all $\gamma < 3\inf\{\beta_1^2, \beta_2^2\} - 1$, the semi-Hölder norm of exponent $\gamma/3$ of $D_{x_2} u_i$ w.r.t $x_2$ is bounded uniformly in $t$ and $x_1$ :*

$$||D_{x_2} u_i||_{\infty,\infty,\gamma/3} := \sup_{w_1, w_2, w_2' \in \mathbb{R}^d,\ t \in [0,T]} \frac{|D_{x_2} u_i(t, w_1, w_2) - D_{x_2} u_i(t, w_1, w_2')|}{|w_2 - w_2'|^{\gamma/3}} \leq CT^{\delta_{1.4.6}},$$

*where $C$ and $\delta_{1.4.6} > 0$ depend only on known parameters in **(H)**.*

*Proof of Claim 1.4.6 :* Let us first introduce the quantity :

$$M(D_{x_2} u_i, T) := \sup_{w_1, w_2, w_2' \in \mathbb{R}^d,\ t \in [0,T]} \frac{|D_{x_2} u_i(t, w_1, w_2) - D_{x_2} u_i(t, w_1, w_2')|}{|w_2 - w_2'|^{\gamma/3} + |w_2 - w_2'|^{\beta_2^2} + |w_2 - w_2'|^{\beta_1^2} + |w_2 - w_2'|}. \tag{1.4.14}$$

From (1.3.12) and Lemma 1.3.4, for all $(t, x_1)$ in $[0, T] \times \mathbb{R}^d$ and $(x_2, z_2)$ in $\mathbb{R}^d \times \mathbb{R}^d$ we have :

$$\begin{aligned}&|D_{x_2} u_i(t, x_1, x_2) - D_{x_2} u_i(t, x_1, z_2)| \\ &\leq \left|\sum_{j=1}^{4} \left(D_{x_2} H_i^j\right)(t+\epsilon; t, x_1, x_2) - \left(D_{x_2} H_i^j\right)(t+\epsilon; t, x_1, z_2)\right| + \mathcal{O}(\epsilon).\end{aligned} \tag{1.4.15}$$

We recall that $(H_i^j,\ j=1,\cdots,4)$ depend on the freezing point $\xi=(\xi_1,\xi_2)$ of the process which started from $(x_1,x_2)$ and $(x_1,z_2)$ at time $t$. Here, we choose the same freezing point "$\xi$" for the two processes (with different initial conditions). Let us note that, from (1.3.12) each $D_{x_2}H_i^j$ can be written as :

$$D_{x_2}H_i^j(t+\epsilon;t,x_1,x_2)=\int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}}F_i^j(s,y_1,y_2,\theta^1_{t,s}(\xi),\theta^2_{t,s}(\xi))D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds,$$

where $F_i^j$ is some function properly defined by (1.3.12). Let us split the time interval w.r.t. the characteristic scale of the second component of the system (1.3.1) : we set $\mathcal{S}=\{s\in(t,T]\text{ s.t. }|x_2-z_2|<(s-t)^{3/2}\}$ and $\mathcal{S}^c=\{s\in(t,T]\text{ s.t. }|x_2-z_2|\geq(s-t)^{3/2}\}$.. We have :

$$\begin{aligned}
&\sum_{j=1}^4\left(D_{x_2}H_i^j\right)(t+\epsilon;t,x_1,x_2)-\left(D_{x_2}H_i^j\right)(t+\epsilon;t,x_1,z_2)\\
&=\sum_{j=1}^4\int_{t+\epsilon}^T\mathbf{1}_{\mathcal{S}}\int_{\mathbb{R}^{2d}}\Big\{F_i^j(s,y_1,y_2,\theta^1_{t,s}(\xi),\theta^2_{t,s}(\xi))\\
&\qquad\Big(D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)-D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\Big)\Big\}dy_1dy_2ds\\
&\quad+\sum_{j=1}^4\int_{t+\epsilon}^T\mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\Big\{F_i^j(s,y_1,y_2,\theta^1_{t,s}(\xi),\theta^2_{t,s}(\xi))\\
&\qquad\Big(D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)-D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\Big)\Big\}dy_1dy_2ds\\
&:=\sum_{j=1}^4P_i^j(t+\epsilon;t,x;z,\mathcal{S})+\sum_{j=1}^4P_i^j(t+\epsilon;t,x;z,\mathcal{S}^c).
\end{aligned}\tag{1.4.16}$$

As a first step, we bound the sum $\sum_{j=1}^4P_i^j(t+\epsilon;t,x;z,\mathcal{S})$ in (1.4.16). We first prove that for all $s$ in $\mathcal{S}$ the following inequality holds :

$$\begin{aligned}
&|D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)-D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)|\\
&\quad\leq C(s-t)^{-(3+\gamma)/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)|x_2-z_2|^{\gamma/3},
\end{aligned}\tag{1.4.17}$$

where $c$ and $C$ depend only on known parameters in **(H)** and where $0<\gamma<3$.

By using Mean Value Theorem and the Gaussian estimate of $D^2_{x_2}\tilde{q}$ from Proposition 1.3.1 we have :

$$\begin{aligned}
&|D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)-D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)|\\
&\quad\leq\sup_{\rho\in(0,1)}\left|D^2_{x_2}\tilde{q}(t,x_1,x_2+\rho(x_2-z_2);s,y_1,y_2)\right||x_2-z_2|\\
&\quad\leq C'(s-t)^{-3}\sup_{\rho\in(0,1)}\hat{q}_{\bar{c}}(t,x_1,x_2+\rho(x_2-z_2);s,y_1,y_2)\,|x_2-z_2|\,,
\end{aligned}\tag{1.4.18}$$

where $\bar{c}$ is a positive constant depending only on known parameters in **(H)**. Note that on $\mathcal{S}$ :

$$\sup_{\rho\in(0,1)}\hat{q}_{\bar{c}}(t,x_1,x_2+\rho(x_2-z_2);s,y_1,y_2)\leq C''\hat{q}_c(t,x_1,x_2;s,y_1,y_2).\tag{1.4.19}$$

Combining (1.4.18) and (1.4.19), we obtain :

$$\begin{aligned}&|D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)-D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)|\\&\quad\leq C'''(s-t)^{-3}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)\,|x_2-z_2|\,.\end{aligned}$$

Rewrite $|x_2-z_2|=|x_2-z_2|^{1-\gamma/3}|x_2-z_2|^{\gamma/3}$. Since $|x_2-z_2|<(s-t)^{3/2}$ we have $|x_2-z_2|<(s-t)^{3/2-\gamma/2}|x_2-z_2|^{\gamma/3}$ and (1.4.17) follows.

*Bound of* $\sum_{j=1}^3\left|P_i^j(t+\epsilon;t,x;z,\mathcal{S})\right|$. Following the proof of Lemma 1.4.2 and using (1.4.17) instead of (1.4.6) when bounding in the terms (1.4.7), (1.4.9) and (1.4.10) when $\xi=x$, we deduce that :

$$\begin{aligned}&\sum_{j=1}^3\left|P_i^j(t+\epsilon;t,x;z,\mathcal{S})\right| \qquad (1.4.20)\\&\leq C\int_{t+\epsilon}^T\Bigg\{(s-t)^{-(3+\gamma-3\beta_i^2)/2}+\left\|D_{x_1}^2u_i\right\|_\infty(s-t)^{-\gamma/2}+\|D_{x_1}D_{x_2}u_i\|_\infty(s-t)^{-\gamma/2}\\&+\|D_{x_1}u_i\|_\infty(s-t)^{-(3+\gamma-3\beta_1^2)/2}+\|D_{x_1}D_{x_2}u_i\|_\infty(s-t)^{-(\gamma-\beta_1^1)/2}\Bigg\}ds|x_2-z_2|^{\gamma/3},\end{aligned}$$

for all $\gamma<3\inf(\beta_1^2,\beta_2^2)-1$.

*Bound of* $P_i^4(t+\epsilon;t,x;z,\mathcal{S})$. Thanks to (1.4.13), we have :

$$\begin{aligned}&P_i^4(t+\epsilon;t,x;z,\mathcal{S})\\&=-\int_{t+\epsilon}^T\mathbf{1}_{\mathcal{S}}\int_{\mathbb{R}^{2d}}\Big\{\left[\Delta^1(\theta_{t,s}(\xi))F_2(s,y_1,\theta_{t,s}^2(\xi))-D_{x_1}F_2(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi))\Delta^1(\theta_{t,s}(\xi))y_1\right]\\&\quad\cdot\left[\Delta^2(\theta_{t,s}(\xi))D_{x_2}u_i(s,y_1,y_2)\right]+\left[\Delta^2(\theta_{t,s}(\xi))F_2(s,y_1,y_2)\right]\cdot D_{x_2}u_i(s,y_1,y_2)\Big\}\\&\quad\times\left[D_{x_2}\tilde{q}(t,x_1,x_2;t,y_1,y_2)-D_{x_2}\tilde{q}(t,x_1,z_2;t,y_1,y_2)\right]dy_1dy_2ds.\end{aligned}$$

From (1.4.2) and (1.4.17), we deduce that :

$$\begin{aligned}&\left|P_i^4(t+\epsilon;t,x;z,\mathcal{S})\right| \qquad (1.4.21)\\&\leq C'\int_{t+\epsilon}^T\mathbf{1}_{\mathcal{S}}\int_{\mathbb{R}^{2d}}\Bigg\{\left\|D_{x_2}u_i\right\|_\infty(s-t)^{3(\beta_2^2-1-\gamma/3)/2}\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|^{\beta_2^2}\\&\quad+M(D_{x_2}u_i,T)(s-t)^{-1-\gamma/2+\alpha^1/2}\left|\frac{\Delta^1(\theta_{t,s}(\xi))y_1}{(s-t)^{1/2}}\right|^{1+\alpha^1}\Bigg((s-t)^{\gamma/2}\left|\frac{y_2-\theta_{t,s}^2(\xi)}{(s-t)^{3/2}}\right|^{\gamma/3}\\&\quad+(s-t)^{3\beta_1^2/2}\left|\frac{y_2-\theta_{t,s}^2(\xi)}{(s-t)^{3/2}}\right|^{\beta_1^2}+(s-t)^{3\beta_2^2/2}\left|\frac{y_2-\theta_{t,s}^2(\xi)}{(s-t)^{3/2}}\right|^{\beta_2^2}\\&\quad+(s-t)^{3/2}\left|\frac{y_2-\theta_{t,s}^2(\xi)}{(s-t)^{3/2}}\right|\Bigg)\Bigg\}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds|x_2-z_2|^{\gamma/3},\end{aligned}$$

for all $\gamma < 3\beta_2^2 - 1$ and where $M(D_{x_2}u_i, T)$ is defined by (1.4.14).

By setting $\xi = x$ in (1.4.21) and letting $\epsilon \to 0$ in (1.4.20) and (1.4.21), we deduce that there exist a positive constants $C$ and a positive number $\delta^1_{1.4.6}$, depending only on known parameters in **(H)**, such that :

$$\left|\sum_{j=1}^{4} P_i^j(t,x;z,\mathcal{S})\right| \leq CT^{\delta^1_{1.4.6}} \left(M(D_{x_2}u_i,T) + ||D_{x_1x_2}u_i||_\infty\right) |x_2 - z_2|^{\gamma/3}, \tag{1.4.22}$$

for all $\gamma < 3\inf(\beta_1^2, \beta_2^2) - 1$.

As a second step, we bound the sum $\sum_{j=1}^4 P_i^j(t+\epsilon;t,x;z,\mathcal{S}^c)$ in (1.4.16). Note that this sum reads :

$$\begin{aligned}
&\sum_{j=1}^{4} P_i^j(t+\epsilon;t,x;z,\mathcal{S}^c)\\
&= \sum_{j=1}^{4}\left\{\int_{t+\epsilon}^{T} \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}} F_i^j(s,y_1,y_2,\theta^1_{t,s}(\xi),\theta^2_{t,s}(\xi)) D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds\right.\\
&\left. - \int_{t+\epsilon}^{T} \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}} F_i^j(s,y_1,y_2,\theta^1_{t,s}(\xi),\theta^2_{t,s}(\xi)) D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)dy_1dy_2ds\right\}\\
&:= \sum_{j=1}^{4}\{\tilde{H}_i^j(t+\epsilon;t,x_1,x_2,\mathcal{S}^c) - \tilde{H}_i^j(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)\},
\end{aligned} \tag{1.4.23}$$

and that for all $s$ in $\mathcal{S}^c$ we have :

$$1 \leq (s-t)^{-\gamma/2}|x_2 - z_2|^{\gamma/3}. \tag{1.4.24}$$

On a first hand, by plugging (1.4.24) in (1.4.7), (1.4.9) and (1.4.11) in the proof of Lemma 1.4.2 we obtain that :

$$\begin{aligned}
&\sum_{j=1}^{3}\left|\tilde{H}_i^j(t+\epsilon;t,x_1,x_2,\mathcal{S}^c)\right| \qquad (1.4.25)\\
&\leq C\int_{t+\epsilon}^{T}\left\{(s-t)^{-3(1+\gamma/3-\beta_1^2)/2} + \left\|D^2_{x_1}u_i\right\|_\infty (s-t)^{-\gamma/2} + \|D_{x_1}D_{x_2}u_i\|_\infty (s-t)^{-\gamma/2}\right.\\
&\left. + \|D_{x_1}u_i\|_\infty (s-t)^{-3(1+\gamma/3-\beta_1^2)/2} + \|D_{x_1}D_{x_2}u_i\|_\infty (s-t)^{-(\gamma-\beta_1^1)/2}\right\} ds|x_2 - z_2|^{\gamma/3}.
\end{aligned}$$

By using (1.4.13) and (1.4.24), and setting $\xi = x$, we deduce that :

$$\begin{aligned}
&\left|\tilde{H}_i^4(t+\epsilon;t,x_1,x_2,\mathcal{S}^c)\right| \qquad (1.4.26)\\
&\leq C\int_{t+\epsilon}^T \Bigg\{ \|D_{x_2}u_i\|_\infty (s-t)^{3(\beta_2^2-1-\gamma/3)/2} + CM(D_{x_2}u_i,T)\\
&\quad \times (s-t)^{-1+\alpha^1/2}\Big((s-t)^{\gamma/2}+(s-t)^{3\beta_1^2/2}+(s-t)^{3\beta_2^2/2}+(s-t)^{3/2}\Big)\Bigg\}ds\\
&\quad \times |x_2-z_2|^{\gamma/3},
\end{aligned}$$

where the quantity $M(D_{x_2}u_i,T)$ is defined by (1.4.14).

On a second hand, we have to deal with the terms $\tilde{H}_i^j(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)$, for $j \in \{1,\cdots,4\}$. Since we take the same freezing point for the two solutions with different initial conditions, we have to re-center each integrand of $\tilde{H}_i^j(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)$ in order to use the off-diagonal decay of the Gaussian exponential w.r.t the degenerate component. In this case, for all $s$ in $[t,T]$ and $y_2$ in $\mathbb{R}^d$, this off-diagonal decay is given by : $|y_2-m_{t,s}^{2,\xi}(x_1,z_2)|$. Hence, we have to center the terms around $(m_{t,s}^{2,\xi}(x_1,z_2))_{t<s\leq T}$ with Claim 1.3.3.

*Bound of* $\tilde{H}_i^1(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)$. We have from Claim 1.3.3, estimate on $D_{x_2}\tilde{q}$ from Proposition 1.3.1 and (1.4.24) that :

$$\begin{aligned}
&\left|\tilde{H}_i^1(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)\right| \qquad (1.4.27)\\
&= \left|\int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}} \mathbf{1}_{\mathcal{S}^c}\Delta^2(m_{t,s}^{2,\xi}(x_1,z_2))\phi_i(s,y_1,y_2)D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)dy_1dy_2ds\right|\\
&\leq C\int_{t+\epsilon}^T\Bigg\{\mathbf{1}_{\mathcal{S}^c}(s-t)^{-3(1-\beta_i^2+\gamma/3)/2}\int_{\mathbb{R}^{2d}}\left|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\right|^{\beta_i^2}\\
&\quad \times \hat{q}_c(t,x_1,z_2;s,y_1,y_2)\Bigg\}dy_1dy_2ds|x_2-z_2|^{\gamma/3},
\end{aligned}$$

for all $\gamma < 3\beta_i^2-1$.

*Bound of* $\tilde{H}_i^2(t+\epsilon; t, x_1, z_2, \mathcal{S}^c)$. By using Claim 1.3.3, an integration by parts (as in the proof of Claim 1.4.4) we obtain :

$$\begin{aligned}
&\tilde{H}_i^2(t+\epsilon; t, x_1, z_2, \mathcal{S}^c)\\
&= \frac{1}{2}\int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\Big\{\mathrm{Tr}\Big[\Big[\Delta^2(m_{t,s}^{2,\xi}(x_1,z_2))a(s,y_1,y_2)\Big]\, D_{y_1}^2 u_i(s,y_1,y_2)\Big]\\
&\qquad \times [D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)]\Big\}dy_1dy_2ds\\
&\quad +\frac{1}{2}\sum_{l=1}^d \int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\Big\{\Big(\Big[\frac{\partial}{\partial y_{1l}} a_{l.}(s,y_1,m_{t,s}^{2,\xi}(x_1,z_2))\Big]\\
&\qquad \cdot\Big[\Delta^2(m_{t,s}^{2,\xi}(x_1,z_2))D_{y_1}u_i(s,y_1,y_2)\Big]\Big)\,[D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)]\Big\}dy_1dy_2ds\\
&\quad +\frac{1}{2}\sum_{l=1}^d \int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\Big\{\Big[a_{l.}(s,y_1,m_{t,s}^{2,\xi}(x_1,z_2)) - a_{l.}(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi))\Big]\\
&\qquad \cdot\Big[\Delta^2(m_{t,s}^{2,\xi}(x_1,z_2))D_{y_1}u_i(s,y_1,y_2)\Big]\times\Big[D_{x_2}\Big(\frac{\partial}{\partial y_{1l}}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\Big)\Big]\Big\}dy_1dy_2ds,
\end{aligned}$$

From (1.4.24), regularity of $a$ from **(H)** and Proposition 1.3.1 we have :

$$\begin{aligned}
&\Big|\tilde{H}_i^2(t+\epsilon; t, x_1, z_2, \mathcal{S}^c)\Big| \qquad (1.4.28)\\
&\le C\int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\Big\{\Big[||D_{x_1^2}u_i||_\infty (s-t)^{-\gamma/2}\Big|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\Big|\\
&\qquad +||D_{x_1}D_{x_2}u_i||_\infty (s-t)^{-\gamma/2}\Big|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\Big|\\
&\qquad +(s-t)^{-\gamma/2}||D_{x_1}D_{x_2}u_i||_\infty\Big|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\Big|\Big|\frac{y_1-\theta_{t,s}^1(\xi)}{(s-t)^{1/2}}\Big|\Big]\Big\}\\
&\qquad\quad \times \hat{q}_c(t,x_1,z_2;s,y_1,y_2)dy_1dy_2ds|x_2-z_2|^{\gamma/3}\\
&\quad +C||D_{x_1}D_{x_2}u_i||_\infty\Big\{\int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}(s-t)^{-1/2}\int_{\mathbb{R}^{2d}}\Big|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\Big|\hat{q}_c(t,x_1,z_2;s,y_1,y_2)dy_1dy_2ds\Big\}\\
&\qquad\quad \times|m_{t,s}^{2,\xi}(x_1,z_2)-\theta_{t,s}^2(\xi)|,
\end{aligned}$$

for all $\gamma < 2$.

*Bound of* $\tilde{H}_i^3(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)$. From Claim 1.3.3 and representation 1.4.4, this term can be centered w.r.t. the coefficients and $D_{x_1}u_i$ as follows :

$$\begin{aligned}
&\tilde{H}_i^3(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)\\
&= \int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}} \left[\Delta^2(m_{t,s}^{2,\xi}(x_1,z_2))F_1(s,y_1,y_2)\right]\\
&\quad\cdot \left[D_{x_1}u_i(s,y_1,y_2)\right] D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)dy_1dy_2ds\\
&\quad + \int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}} \left[F_1(s,y_1,m_{t,s}^{2,\xi}(x_1,z_2)) - F_1(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi))\right]\\
&\quad\cdot \left[\Delta^2(m_{t,s}^{2,\xi}(x_1,z_2)D_{x_1}u_i(s,y_1,y_2))\right] D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)dy_1dy_2ds.
\end{aligned}$$

From (1.4.24), (1.4.4) and Proposition 1.3.1, we can deduce the following :

$$\begin{aligned}
&\left|\tilde{H}_i^3(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)\right| \qquad (1.4.29)\\
&\le C\int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}} \Bigg\{\Bigg[||D_{x_1}u_i||_\infty (s-t)^{-3(1-\beta_1^2+\gamma/3)/2}\left|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\right|^{\beta_1^2}\\
&\quad + ||D_{x_1}D_{x_2}u_i||_\infty (s-t)^{(\beta_1^1-\gamma)/2}\left|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\right|\left|\frac{y_1-\theta_{t,s}^1(\xi)}{(s-t)^{1/2}}\right|^{\beta_1^1}\Bigg]\\
&\quad\times \hat{q}_c(t,x_1,z_2;s,y_1,y_2)\Bigg\}dy_1dy_2ds|x_2-z_2|^{\gamma/3}\\
&+C||D_{x_1}D_{x_2}u_i||_\infty\left\{\int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}\int_{\mathbb{R}^{2d}}\left|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\right|\hat{q}_c(t,x_1,z_2;s,y_1,y_2)dy_1dy_2ds\right\}\\
&\quad\times|m_{t,s}^{2,\xi}(x_1,z_2)-\theta_{t,s}^2(\xi)|^{\beta_1^2}
\end{aligned}$$

for all $\gamma < 3\beta_1^2 - 1$.

*Bound of* $\tilde{H}_i^4(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)$. From Claim 1.3.3, by centering w.r.t. $D_{x_2}u_i$, this term can be written as :

$$\begin{aligned}
&\tilde{H}_i^4(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)\\
&= \int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}}\Big\{\mathbf{1}_{\mathcal{S}^c}\left[\Delta^2(m_{t,s}^{2,\xi}(x_1,z_2))F_2(s,y_1,y_2)\right]\cdot\left[D_{x_2}u_i(s,y_1,y_2)\right]\\
&\quad\times D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)\Big\}dy_1dy_2ds\\
&\quad+\int_{t+\epsilon}^T\int_{\mathbb{R}^{2d}}\Big\{\mathbf{1}_{\mathcal{S}^c}\Big[F_2(s,y_1,m_{t,s}^{2,\xi}(x_1,z_2)) - F_2(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi))\\
&\quad - D_{x_1}F_2(s,\theta_{t,s}^1(\xi),\theta_{t,s}^2(\xi))\Delta^1(\theta_{t,s}(\xi))y_1\Big]\cdot\left[\Delta^2(m_{t,s}^{2,\xi}(x_1,z_2))D_{x_2}u_i(s,y_1,y_2)\right]\Big\}\\
&\quad\times D_{x_2}\tilde{q}(t,x_1,z_2;s,y_1,y_2)dy_1dy_2ds.
\end{aligned}$$

By using regularity of the coefficients from **(H)**, Proposition 1.3.1 and (1.4.24) we have :

$$
\begin{aligned}
&\left|\tilde{H}_i^4(t+\epsilon;t,x_1,z_2,\mathcal{S}^c)\right| \qquad (1.4.30)\\
&\le C||D_{x_2}u_i||_\infty \int_{t+\epsilon}^T \Bigg\{ \mathbf{1}_{\mathcal{S}^c}(s-t)^{-3(1-\beta_2^2+\gamma/3)/2}\\
&\quad \times \int_{\mathbb{R}^{2d}} \left|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\right|^{\beta_2^2} \hat{q}_c(t,x_1,z_2;s,y_1,y_2)dy_1dy_2\Bigg\} ds|x_2-z_2|^{\gamma/3}\\
&\quad + CM(D_{x_2}u_i,T)\int_{t+\epsilon}^T \mathbf{1}_{\mathcal{S}^c}(s-t)^{-1+\alpha^1/2}\int_{\mathbb{R}^{2d}} \Bigg\{ \left|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\right|^{\gamma/3} |x_2-z_2|^{\gamma/3}\\
&\quad + \left|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\right|^{\beta_1^2} |x_2-z_2|^{\beta_1^2} + \left|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\right|^{\beta_2^2} |x_2-z_2|^{\beta_2^2}\\
&\quad + \left|\frac{y_2-m_{t,s}^{2,\xi}(x_1,z_2)}{(s-t)^{3/2}}\right| |x_2-z_2|\Bigg\} \hat{q}_c(t,x_1,z_2;s,y_1,y_2)dy_1dy_2ds,
\end{aligned}
$$

for all $\gamma < 3\beta_1^2/2 - 1$.

Now, note that from (1.3.3),

$$m_{t,s}^{2,x}(x_1,z_2) - \theta_{t,s}^2(x) = z_2 - x_2.$$

Hence, by setting $\xi = x$ and letting $\epsilon \to 0$ in (1.4.27), (1.4.28), (1.4.29) and (1.4.30) and combining the resulting estimates with (1.4.25) and (1.4.26), we deduce that there exist a positive constant $C$ and a positive number $\delta_{1.4.6}^2$ depending only on known parameters in **(H)**, such that :

$$
\begin{aligned}
\left|\sum_{j=1}^4 P_i^j(t,x;z,\mathcal{S}^c)\right| \quad &\le \quad CT^{\delta_{1.4.6}^2}\left(M(D_{x_2}u_i,T)+||D_{x_1}D_{x_2}u_i||_\infty\right) \qquad (1.4.31)\\
&\qquad \times \left(|x_2-z_2|^{\gamma/3}+|x_2-z_2|^{\beta_2^2}+|x_2-z_2|^{\beta_1^2}+|x_2-z_2|\right),
\end{aligned}
$$

for all $\gamma < 3\inf(\beta_1^2,\beta_2^2)-1$.

Finally, by plugging estimates (1.4.22) and (1.4.31) in (1.4.15), we deduce that there exist a positive constant $C$ and a positive number $\delta_{1.4.6}$ such that :

$$
\begin{aligned}
&|D_{x_2}u_i(t,x_1,x_2) - D_{x_2}u_i(t,x_1,z_2)|\\
&\le CT^{\delta_{1.4.6}}\left(1+C_T'M(D_{x_2}u_i,T)\right)\left(|x_2-z_2|^{\gamma/3}+|x_2-z_2|^{\beta_2^2}+|x_2-z_2|^{\beta_1^2}+|x_2-z_2|\right).
\end{aligned}
$$

Together with the boundedness of $D_{x_2}u_i$ from Lemma 1.4.2, this concludes the proof of Claim 1.4.6. □

Now, we prove of Lemma 1.4.3. It follows from the representation (1.3.12) and Lemma

1.3.4 that, for all $\epsilon > 0$ :

$$D_{x_1} D_{x_2} u_i(t, x_1, x_2) = \sum_{j=1}^{4} D_{x_1} D_{x_2} H_i^j(t+\epsilon; t, x_1, x_2) + \mathcal{O}(\epsilon) \tag{1.4.32}$$

We bound each $D_{x_1} D_{x_2} H_i^j$, $j = 1, \ldots, 4$. We recall that from Proposition 1.3.1, there exists a positive constant $C$ depending only on known parameters in **(H)** such that :

$$D_{x_1} D_{x_2} \tilde{q}(t, x_1, x_2; s, y_1, y_2) \leq C(s-t)^{-2} D_{x_1} D_{x_2} \hat{q}_c(t, x_1, x_2; s, y_1, y_2). \tag{1.4.33}$$

*Bound of $D_{x_1} D_{x_2} H_i^1$.* From Claim 1.3.3 and (1.4.33) we have :

$$\begin{aligned}
&\left| D_{x_1} D_{x_2} H_i^1(t+\epsilon; t, x_1, x_2) \right| \\
&= \left| \int_{t+\epsilon}^{T} \int_{\mathbb{R}^{2d}} \Delta^2(\theta_{t,s}(\xi)) \phi_i(s, y_1, y_2) \left[ D_{x_1} D_{x_2} \tilde{q}(t, x_1, x_2; s, y_1, y_2) \right] dy_1 dy_2 ds \right| \\
&\leq C' \int_{t+\epsilon}^{T} \int_{\mathbb{R}^{2d}} (s-t)^{-2+3/2\beta_i^2} \left| \frac{\Delta^2(\theta_{t,s}(\xi)) y_2}{(s-t)^{3/2}} \right|^{\beta_i^2} \hat{q}_c(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds.
\end{aligned}$$

Since $\beta_i^2 > 2/3,\ i = 1, 2$ from **(H1)**, by setting $\xi = x$ and letting $\epsilon \to 0$ we deduce that :

$$\left\| D_{x_1} D_{x_2} H_i^1 \right\|_\infty \leq C'' T^{(3\beta_i^2/2-1)}. \tag{1.4.34}$$

*Bound of $D_{x_1} D_{x_2} H_i^2$.* Thanks to (1.4.3), we have

$$\begin{aligned}
&D_{x_1} D_{x_2} H_i^2(t+\epsilon; t, x_1, x_2) \\
&= -\frac{1}{2} \int_{t+\epsilon}^{T} \int_{\mathbb{R}^{2d}} \Big\{ \mathrm{Tr}\left[ [\Delta^2(\theta_{t,s}(\xi)) a(s, y_1, y_2)] D_{x_1}^2 u_i(s, y_1, y_2) \right] \\
&\quad \times \left[ D_{x_1} D_{x_2} \tilde{q}(t, x_1, x_2; s, y, y_1, y_2) \right] \Big\} dy_1 dy_2 ds \\
&\quad + \frac{1}{2} \sum_{l=1}^{d} \int_{t+\epsilon}^{T} \int_{\mathbb{R}^{2d}} \left\{ \left[ \frac{\partial}{\partial y_{1l}} a_{l.}(s, y_1, \theta_{t,s}^2(\xi)) \right] \cdot \left[ \Delta^2(\theta_{t,s}(\xi)) D_{x_1} u_i(s, y_1, y_2) \right] \right. \\
&\quad \left. \times \left[ D_{x_1} D_{x_2} \tilde{q}(t, x_1, x_2; s, y_1, y_2) \right] \right\} dy_1 dy_2 ds \\
&\quad + \frac{1}{2} \sum_{l=1}^{d} \int_{t+\epsilon}^{T} \int_{\mathbb{R}^{2d}} \Big\{ \left[ \Delta^1(\theta_{t,s}(\xi)) a_{l.}(s, y_1, \theta_{t,s}^2(\xi)) \right] \\
&\quad \cdot \left[ \Delta^2(\theta_{t,s}(\xi)) D_{x_1} u_i(s, y_1, y_2) \right] \\
&\quad \times \left[ D_{x_1} D_{x_2} \left( \frac{\partial}{\partial y_{1l}} \tilde{q}(t, x_1, x_2; s, y_1, y_2) \right) \right] \Big\} dy_1 dy_2 ds.
\end{aligned}$$

By using (1.4.33) and regularity of $a$ from **(H)**, we deduce that :

$$\begin{aligned}
&\left|D_{x_1}D_{x_2}H_i^2(t+\epsilon;t,x_1,x_2)\right| \\
&\leq C'\int_{t+\epsilon}^{T}(s-t)^{-1/2}\int_{\mathbb{R}^{2d}}\left\{\left\|D_{x_1}^2u_i\right\|_\infty\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|+\|D_{x_1}D_{x_2}u_i\|_\infty\right. \\
&\quad\left.\left(\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|+\left|\frac{\Delta^1(\theta_{t,s}(\xi))y_1}{(s-t)^{1/2}}\right|\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|\right)\right\}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds.
\end{aligned}$$

By setting $\xi = x$ and letting $\epsilon \to 0$ we obtain :

$$\left\|D_{x_1}D_{x_2}H_i^2\right\|_\infty \leq C''T^{1/2}\left(\|D_{x_1}D_{x_2}u_i\|_\infty+\left\|D_{x_1}^2u_i\right\|_\infty\right). \tag{1.4.35}$$

*Bound of* $D_{x_1}D_{x_2}H_i^3$. From (1.4.4) :

$$\begin{aligned}
&D_{x_1}D_{x_2}H_i^3(t+\epsilon;t,x_1,x_2) \\
&=-\int_{t+\epsilon}^{T}\int_{\mathbb{R}^{2d}}\left\{\left[\Delta^2(\theta_{t,s}(\xi))F_1(s,y_1,y_2)\right]\cdot D_{x_1}u_i(s,y_1,y_2)\right. \\
&\qquad\left.\times\left[D_{x_1}D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right]\right\}dy_1dy_2ds \\
&\quad-\int_{t+\epsilon}^{T}\int_{\mathbb{R}^{2d}}\left\{\left[\Delta^1(\theta_{t,s}(\xi))F_1(s,y_1,\theta_{t,s}^2(\xi))\right]\cdot\left[\Delta^2(\theta_{t,s}(\xi))D_{x_1}u_i(s,y_1,y_2)\right]\right. \\
&\qquad\left.\times\left[D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2)\right]\right\}dy_1dy_2ds.
\end{aligned}$$

So that, by (1.4.33), regularity of $F_1$ from **(H1)** and Mean Value Theorem :

$$\begin{aligned}
&\left|D_{x_1}D_{x_2}H_i^3(t+\epsilon;t,x_1,x_2)\right| \\
&\leq C'\left\{\|D_{x_1}u_i\|_\infty\int_{t+\epsilon}^{T}(s-t)^{-2+3\beta_1^2/2}\int_{\mathbb{R}^{2d}}\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|^{\beta_1^2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)dy_1dy_2ds\right. \\
&\quad+\|D_{x_1}D_{x_2}u_i\|_\infty\int_{t+\epsilon}^{T}(s-t)^{(\beta_1^1-1)/2}\int_{\mathbb{R}^{2d}}\left\{\left|\frac{\Delta^1(\theta_{t,s}(\xi))y_1}{(s-t)^{1/2}}\right|^{\beta_1^1}\left|\frac{\Delta^2(\theta_{t,s}(\xi))y_2}{(s-t)^{3/2}}\right|\right. \\
&\qquad\left.\left.\times\hat{q}_c(t,x_1,x_2;s,y_1,y_2)\right\}dy_1dy_2ds\right\}.
\end{aligned}$$

By setting $\xi = x$ and letting $\epsilon \to 0$ we obtain the following bound :

$$\left\|D_{x_1}D_{x_2}H_i^3\right\|_\infty \leq C''\left(T^{3\beta_1^2/2-1}+T^{(1+\beta_1^1)/2}\right)\|D_{x_1}D_{x_2}u_i\|_\infty. \tag{1.4.36}$$

*Bound of* $D_{x_1}D_{x_2}H_i^4$. Thanks to (1.4.13), we have

$$
\begin{aligned}
&D_{x_1} D_{x_2} H_i^4(t+\epsilon; t, x_1, x_2)\\
&= -\int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \left[\Delta^1(\theta_{t,s}(\xi)) F_2(s, y_1, \theta^2_{t,s}(\xi)) - D_{x_1} F_2(s, \theta^1_{t,s}(\xi), \theta^2_{t,s}(\xi)) \Delta^1(\theta_{t,s}(\xi)) y_1\right]\\
&\quad \cdot [\Delta^2(\theta_{t,s}(\xi)) D_{x_2} u_i(s, y_1, y_2)] D_{x_1} D_{x_2} \tilde{q}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds\\
&\quad - \int_{t+\epsilon}^T \int_{\mathbb{R}^{2d}} \left[\Delta^2(\theta_{t,s}(\xi)) F_2(s, y_1, y_2)\right] \cdot D_{x_2} u_i(s, y_1, y_2) D_{x_1} D_{x_2} \tilde{q}(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds.
\end{aligned}
$$

By using Claim 1.4.6, the regularity of $F_2$, (1.4.2) and (1.4.33), we obtain :

$$
\begin{aligned}
&\left|D_{x_1} D_{x_2} H_i^4(t+\epsilon; t, x_1, x_2)\right|\\
&\leq C' \left\|D_{x_2} u_i\right\|_{\infty,\infty,\gamma/3} \int_{t+\epsilon}^T (s-t)^{(-3+\gamma+\alpha^1)/2} \int_{\mathbb{R}^{2d}} \left\{ \left| \frac{\Delta^2(\theta_{t,s}(\xi)) y_2}{(s-t)^{3/2}} \right|^{\gamma/3} \right.\\
&\quad \left. \times \hat{q}_c(t, x_1, x_2; s, y_1, y_2) \right\} dy_1 dy_2 ds\\
&\quad + C' \left\|D_{x_2} u_i\right\|_\infty \int_{t+\epsilon}^T (s-t)^{-2+3\beta_2^2/2} \int_{\mathbb{R}^{2d}} \left| \frac{\Delta^2(\theta_{t,s}(\xi)) y_2}{(s-t)^{3/2}} \right|^{\beta_2^2} \hat{q}_c(t, x_1, x_2; s, y_1, y_2) dy_1 dy_2 ds.
\end{aligned}
$$

Since this inequality holds for all $\gamma < 3\inf(\beta_2^2, \beta_1^2) - 1$, $\gamma$ can be chosen such that the first term in the right hand side is integrable [31]. Then, by letting $\xi = x$ and letting $\epsilon \to 0$ we deduce that :

$$
\left\|D_{x_1} D_{x_2} H_i^4\right\|_\infty \leq C'' \left( T^{(3\beta_2^2/2-1} + T^{(\alpha^1+\gamma-1)/2} \right) \left\|D_{x_2} u_i\right\|_\infty. \tag{1.4.37}
$$

Combining (1.4.34), (1.4.35), (1.4.36), (1.4.37) and using estimates on $\|D_{x_1} u_i\|_\infty$, $\left\|D^2_{x_1} u_i\right\|_\infty$ and $\|D_{x_2} u_i\|_\infty$ given in Lemma 1.4.1 and Lemma 1.4.2, estimate of $\|D_{x_2} u_i\|_{\infty,\infty,\gamma/3}$ from Claim 1.4.6, we deduce that there exists a positive number $\bar{\delta}_{1.4.3}$ depending only on known parameters in **(H)** such that, for $T$ small enough :

$$
\|D_{x_1} D_{x_2} u_i\|_\infty \leq C T^{\bar{\delta}_{1.4.3}}.
$$

This concludes the proof of Lemma 1.4.3. □

31. Since $\beta_i^2$ is supposed to be strictly greater than $2/3$ in **(H1)**.

# ANNEXE A

## Differentiability of the flow of the solution of stochastic differential equation with Hölder drift and degenerate noise

### *A.1* Main result

For $T > 0$ we recall the following system introduced in the previous chapter (see (1.1.1)) :

$$\begin{aligned} dX_s^{1,t,x} &= F_1(s, X_s^{1,t,x}, X_s^{2,t,x})ds + \sigma(s, X_s^{1,t,x}, X_s^{2,t,x})dW_s, \quad X_t^{1,t,x} = x_1, \\ dX_s^{2,t,x} &= F_2(s, X_s^{1,t,x}, X_s^{2,t,x})ds, \quad X_t^{2,t,x} = x_2, \end{aligned} \tag{A.1.1}$$

for any $t < s$ in $[0,T]^2$ where $(W_t, t \geq 0)$ is a standard $d$-dimensional Brownian motion defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$. Theorem 1.1.1 says that under assumption **(H)** this system admits a unique strong solution. The objective of this part consists in proving that under the additional assumption :

**(R):** $\forall t \in [0,T],\ \sigma(t,\cdot,\cdot) : (x_1,x_2) \in \mathbb{R}^{2d} \mapsto \sigma(t,x_1,x_2)$ is continuously differentiable with $\gamma_\sigma$-Hölder continuous derivative, $\gamma_\sigma \in (0,1)$,

this system defines a stochastic flow which is a.s. differentiable with Hölder continuous derivative.

**Assumption.**

**(H'):** We say that assumption **(H')** holds if assumption **(H)** holds and assumption **(R)** holds.

**Notations.** We keep the notations introduced in Chapter 1. In addition, for any function $\phi$ from $[0,T] \times \mathbb{R}^d \times \mathbb{R}^d$, any $\alpha, \beta, \gamma \in (0,1)$ we denote by $||\phi||_{\alpha,\cdot,\cdot}$ (resp. $||\phi||_{\cdot,\beta,\cdot}$ and $||\phi||_{\cdot,\cdot,\gamma}$) the Hölder semi-norm with respect to the first (resp. second and third) variable. Again, we denote by $C$, $C'$, $c$, $c'$, etc... some positive constants depending only on known parameters in **(H')**. When we add the dependence with respect to $T$ to a constant, we mean that this constant decreases with $T$ and tends to 0 as $T$ tends to 0. The constants may change from line to line.

We have the following

**Proposition A.1.1.** *Suppose that assumption **(H')** holds. For any $t$ in $[0,T]$ and $x = (x_1,x_2)$ in $\mathbb{R}^{2d}$, the unique strong solution $(X_s^{t,x})_{t\leq s\leq T} = (X_s^{1,t,x}, X_s^{2,t,x})_{t\leq s\leq T}$ of (A.1.1) started from $x$ at time $t$ defines a stochastic flow of which is a.s. differentiable with Hölder continuous derivative.*

## *A.2* Proof of Proposition A.1.1

As explained in Chapter 1, the main trick consists in replacing the drift by the solution of the system of PDE (1.1.6) :

$$\begin{aligned}&\partial_t u_i(t,x_1,x_2)+\mathcal{L}u_i(t,x_1,x_2)=F_i(t,x_1,x_2),\ \text{for } (t,x_1,x_2)\in[0,T)\times\mathbb{R}^d\times\mathbb{R}^d\\&u_i(T,x_1,x_2)=0_{\mathbb{R}^d},\quad i=1,2,\end{aligned}\tag{A.2.1}$$

where $\mathcal{L}$ is given by, for all $\psi$ in $C^{1,2,1}([0,T]\times\mathbb{R}^d\times\mathbb{R}^d,\mathbb{R}^d)$ :

$$\begin{aligned}\mathcal{L}\psi(t,x_1,x_2) \;&=\; \frac{1}{2}Tr(a(t,x_1,x_2))D^2_{x_1}\psi(t,x_1,x_2))+[F_1(t,x_1,x_2)]\cdot[D_{x_1}\psi(t,x_1,x_2)]\\&\quad+[F_2(t,x_1,x_2)]\cdot[D_{x_2}\psi(t,x_1,x_2)]\,.\end{aligned}\tag{A.2.2}$$

This solution satisfies :

**Proposition A.2.1.** *Suppose that assumption **(H')** holds. Then, for $T$ small enough, the PDE (A.2.1) admits a unique classical solution $u=(u_1,u_2)^*$ which is in $C^{1,2,1}([0,T]\times\mathbb{R}^d\times\mathbb{R}^d,\mathbb{R}^d)$.*

*Moreover, there exist a positive constant $C_T$ and a positive number $\gamma_{A.2.1}$, depending only on known parameters in **(H)**, such that :*

$$\begin{aligned}||D_{x_1}u||_\infty \;&+\; ||D_{x_2}u||_\infty+||D^2_{x_1}u||_\infty+||D_{x_2}u||_{\cdot,\cdot,\gamma_{A.2.1}}\\&+||D^2_{x_1}u||_{\cdot,\gamma_{A.2.1},\cdot}+||D^2_{x_1}u||_{\cdot,\cdot,\gamma_{A.2.1}}\le C_T.\end{aligned}\tag{A.2.3}$$

*Besides for every $t$ in $[0,T]$ and for all $x_1$ in $\mathbb{R}^d$ the mapping $D_{x_1}u(t,x_1,\cdot):x_2\in\mathbb{R}^d\mapsto D_{x_1}u(t,x_1,x_2)$ is differentiable and there exists a positive constant $C_T$ such that :*

$$||D_{x_2}D_{x_1}u||_{\cdot,\gamma_{A.2.1},\cdot}+||D_{x_2}D_{x_1}u||_{\cdot,\cdot,\gamma_{A.2.1}}+||D_{x_2}D_{x_1}u||_\infty\le C_T.$$

Then, by using this result and Itô's Formula, we can write the solution of (A.1.1) as :

$$X^{t,x}_s=x-u(t,x)+u(t,X^{t,x}_s)-\int_t^s[D_xu-\mathbf{1}]\,B\sigma(r,X^{t,x}_r)dW_r,\tag{A.2.4}$$

where $B=(\mathrm{Id}_{\mathbb{R}^d\times\mathbb{R}^d},0_{\mathbb{R}^d\times\mathbb{R}^d})^*$ and "$\mathbf{1}$" is defined by (1.1.8).

For every $t\in[0,T]$, let us define the mapping $\phi_t:x\in\mathbb{R}^{2d}\mapsto x-u(t,x)$. At the end of this section, we prove that :

**Lemma A.2.2.** *Suppose that assumption **(H')** holds. Then, $T$ small enough and for every $t$ in $[0,T]$, the mapping $\phi_t:x\mapsto\phi_t(x)$ is uniformly (in time) bounded and is a $C^1$-diffeomorphism whose derivative is uniformly bounded. Moreover, for every $t\in[0,T]$, the mapping $x\mapsto D_x\phi_t^{-1}$ is $\gamma_{A.2.1}$-Hölder continuous.*

Now, set $Y^{t,\phi(x)}_s=\phi_s(X^{t,x}_s)$. Then $Y^{t,\phi(x)}_s$ solves :

$$Y^{t,\phi(x)}_s=\phi_t(x)-\int_t^sA(r,Y^{t,\phi(x)}_s)dW_r,\tag{A.2.5}$$

where for all $s\in[0,T]$ and $y\in\mathbb{R}^{2d}$ :

$$A(s,y)=[D_xu-\mathbf{1}]\,B\sigma(s,\phi_s^{-1}(y)).$$

In order to prove Proposition A.1.1, we only have to show that $Y^{t,y}_s$ defines a stochastic flow which is a.s. differentiable with Hölder continuous derivative.

It is well seen, from Lemma A.2.2, assumption **(H')** and Proposition A.2.1 that for all $s \in [0,T]$, the mapping $y \mapsto A(s,y)$ is continuously differentiable with $\gamma_A$-Hölder continuous derivative, for some $0 < \gamma_A < [\gamma_{A.2.1}^2 \wedge \gamma_\sigma \gamma_{A.2.1}]$.

Therefore, we deduce from classical theory of stochastic flow (see Chapter 2 of [HDK+84]) that $Y_s^{t,x}$ a stochastic flow which is a.s. differentiable with $\gamma$-Hölder continuous derivative for any $0 < \gamma < \gamma_A$. This concludes the proof. □

Note that the Jacobian $\eta^{t,x}$ of the $X^{t,x}$ satisfies :

$$\eta_s^{t,x} = \mathrm{Id}_{2\mathrm{d}} - \sum_{i=1}^{d} \int_t^s D_x A_i(r, \phi_r^{-1}(X_r^{t,x}))\eta_r^{t,x} dW_r, \tag{A.2.6}$$

where $\mathrm{Id}_{2\mathrm{d}}$ is the identity matrix of $\mathcal{M}_{2d}(\mathbb{R})$.

**Proof of Lemma A.2.2**. We first show the diffeomorphism property. From Hadamard's Theorem, this is equivalent to show that for every $t \in [0,T]$,

(i) $\lim_{|x| \to +\infty} |\phi_t(x)| = +\infty$,

(ii) the Jacobian of $\phi_t$ is invertible.

Assertion (i) follows from the definition of $\phi_t$ and estimate (A.2.3) of $||D_x u||_\infty$ from Proposition A.2.1 with $T$ small enough. Next, the function $\phi_t$ is clearly continuously differentiable and for all $x \in \mathbb{R}^{2d}$, $D_x\phi_t(x) = \mathrm{Id}_{2d} - D_x u(t,x)$. Since from Proposition A.2.1 we know, that for $T$ small enough, the series $\sum_{k\geq 0}(D_x u(t,x))^k$ converges, we deduce that $D_x\phi_t(x)$ is invertible and (ii) is proved.

We now prove the boundedness and regularity assumptions. Remark first that, for $T$ small enough, for all $t \in [0,T]$ :

$$\begin{aligned}
\left|D_x\phi_t^{-1}(x)\right| = \left|[D_x\phi_t(\phi_t^{-1}(x))]^{-1}\right| &= \left|[\mathrm{Id}_{2d} - D_x u(t, \phi_t^{-1}(x))]^{-1}\right| \\
&= \left|\sum_{k\geq 0}(D_x u(t, \phi_t^{-1}(x)))^k\right| \\
&\leq \sum_{k\geq 0} \|D_x u\|_\infty^k < +\infty,
\end{aligned}$$

thanks to Proposition A.2.1. The $\gamma_{A.2.2}$-Hölder regularity of $D_x\phi_t^{-1}$ follows from the estimate (A.2.3) on the supremum norm of $D_x u$ and its regularity from Proposition A.2.1. □

## *A.3* Proof of Proposition A.2.1

We prove the solvability of (A.2.1) by a compactness argument. By denoting by $u^n$ the solution of the PDE (A.2.1) with mollified coefficients $(a^n, F_1^n, F_2^n)_{n\geq 0}$ defined by (1.1.7) in the previous chapter, we show the following *a priori* uniform (w.r.t. the regularization procedure) controls :

**Claim A.3.1.** *For $T$ small enough there exists a positive constant $C_T$, depending only on known parameters in **(H)** and $T$, such that :*

$$||D_{x_1}u^n||_\infty + ||D_{x_2}u^n||_\infty + ||D_{x_1}^2 u^n||_\infty + ||D_{x_2}D_{x_1}u^n||_\infty + ||D_{x_2}u^n||_{.,.,\gamma_{A.2.1}} \leq C_T, \tag{A.3.1}$$

*and*
$$||D_{x_k}D_{x_1}u^n||_{\cdot,\cdot,\gamma_{A.2.1}} + ||D_{x_k}D_{x_1}u^n||_{\cdot,\gamma_{A.2.1},\cdot} \le C_T, \tag{A.3.2}$$
*for* $k = 1, 2$.

*Moreover, on every compact subset* $\mathcal{K}$ *of* $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$*, there exists a constant* $K$ *depending only on known parameters in **(H)** and on the supremum norm* $||F||_{\infty,\mathcal{K}}$ *of* $F = (F_1, F_2)^*$ *on* $\mathcal{K}$*, such that*
$$||D_{x_2}u^n||_{\gamma_{A.2.1},\cdot,\cdot} + ||D^2_{x_1}u^n||_{\gamma_{A.2.1},\cdot,\cdot} + ||D_{x_1}u^n||_{\gamma_{A.2.1},\cdot,\cdot} \le K. \tag{A.3.3}$$

Then, thanks to Arzelà-Ascoli Theorem, we can extract a converging subsequence of $u^n$ (*resp.* $D_{x_1}u^n, D_{x_2}u^n$ $D^2_{x_1}u^n$ and $\partial_t u^n$) that converges uniformly in supremum norm to a limit function $u$ (*resp.* $D_{x_1}u, D_{x_2}u$, $D^2_{x_1}u$ and $\partial_t u$) on every compact subset of $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$. Moreover, for every $t$ in $[0,T]$, we can extract a converging subsequence of $(D_{x_2}D_{x_1}u(t,\cdot))_{n\ge 0}$ to $D_{x_2}D_{x_1}u(t,\cdot)$ in supremum norm on every compact subset of $\mathbb{R}^d\times\mathbb{R}^d$. It is then clear that such a function $u$ satisfies PDE (A.2.1) and estimates of Proposition A.2.1. □

**Proof of Claim A.3.1.** For this proof, we forgive the superscript "$n$" that follows from the regularization procedure and prove the result under **(HR)** but our estimates are obtained only in term of known parameters in **(H)**. Note that the estimates (A.3.1) are given by Proposition 1.1.3 in Chapter 1.

We first prove the estimates (A.3.2).

• *Bound of* $||D_{x_k}D_{x_1}u||_{\cdot,\cdot,\gamma_{A.2.1}}$. We can prove this assertion by following the proof of Claim 1.4.6 in the previous chapter. For the case $k = 2$, we then obtain the same estimates as (1.4.22) and (1.4.31) where the bound for the Hölder exponent is $3\inf(\beta_1^2, \beta_2^2) - 2$.

For $k = 1$, we again follow the proof, but with centering arguments of the proof of Lemma 1.4.1. This gives the estimate for any Hölder exponent less than $\inf(\beta_1^1, \beta_2^1, \alpha_1)$.

• *Bound of* $||D_{x_k}D_{x_1}u||_{\cdot,\gamma_{A.2.1},\cdot}$. Again, this can be proved by following the proof of Claim 1.4.6 in the previous chapter. Instead of (1.4.17), we have that for all $\lambda\in(0,1)$ and for all $s\in(t,T]$ such that $|x_1 - z_1| < (s-t)^{1/2}$,
$$\begin{aligned}&\big|D^2_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2) - D^2_{x_1}\tilde{q}(t,z_1,x_2;s,y_1,y_2)\big|\\ &\quad\le C(s-t)^{-(2+\gamma)/2}\hat{q}_c(t,x_1,x_2;s,y_1,y_2)|x_1-z_1|^{\gamma}.\end{aligned} \tag{A.3.4}$$
Indeed, in that case,
$$\hat{q}_c(t,\lambda x_1 + (1-\lambda)z_1, x_2; s, y_1, y_2) \le \text{const}.\hat{q}_{c'}(t,x_1,x_2;s,y_1,y_2), \tag{A.3.5}$$
for $c' < c$ and where
$$\text{const.} \propto \exp\left(\int_t^s D_{x_1}F_2(r,\theta_{t,r}(\xi))dr\frac{|x_1-z_1|}{(s-t)^{3/2}}\right) \le C\exp(||D_{x_1}F_2||_\infty).$$
Hence, when $k = 2$, we follow the proof of Claim 1.4.6 in the previous chapter and we obtain the same estimates as (1.4.22) and (1.4.31) where the bound for the Hölder exponent is $3\inf(\beta_1^2,\beta_2^2)-2$. When $k=1$ we also follow the proof of Claim 1.4.6 in the spirit

of the proof of Lemma 1.4.1. This gives the estimate for any Hölder exponent less than $\inf(\beta_1^1, \beta_2^1, \alpha_1)$.

We now prove the estimates (A.3.3). Let $\mathcal{K}$ be a compact subset of $[0,T] \times \mathbb{R}^d \times \mathbb{R}^d$.

$\bullet$*Bound of* $||D_{x_2} u||_{\gamma_{A.2.1},.,.}$. Let $\tilde{u} = (\tilde{u}_1, \tilde{u}_2)$ be the solution of the PDE (A.2.1) with the boundary condition $u(s,.)$ at time $s \in (t,T)$. Let $i \in \{1,2\}$, according to (1.3.12), the solution $\tilde{u}_i$ writes :

$$\tilde{u}_i^n(t,x) = \int_t^s \mathbb{E}\left[\phi_i(r, \tilde{X}_r^{t,x}) - (\mathcal{L} - \tilde{\mathcal{L}}^{t,\xi})\tilde{u}_i(r, \tilde{X}_r^{t,x})\right] dr + \int_{\mathbb{R}^{2d}} u_i(s,y)\tilde{q}(t,x;s,y)dy,$$

where $\tilde{X}^{t,x} = (\tilde{X}^{2,t,x}, \tilde{X}^{1,t,x})^*$ is the solution of the frozen system (1.3.1) with initial condition $x$ at time $t$, where $\tilde{\mathcal{L}}^{t,\xi}$ is given by (1.3.6) and where $\tilde{q}$ is defined by (1.3.5) in Proposition 1.3.1. Then,

$$\begin{aligned} D_{x_2}\tilde{u}_i(t,x) &= D_{x_2}\int_t^s \mathbb{E}\left[\phi_i(r, \tilde{X}_r^{t,x}) - (\mathcal{L} - \tilde{\mathcal{L}}^{t,\xi})\tilde{u}_i(r, \tilde{X}_r^{t,x})\right] dr \\ &\quad + \int_{\mathbb{R}^{2d}} u_i(s,y) D_{x_2}\tilde{q}(t,x;s,y)dy. \end{aligned}$$

Since the proof of Proposition 1.3.1 shows that

$$D_{x_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2) = -D_{y_2}\tilde{q}(t,x_1,x_2;s,y_1,y_2),$$

an integration by parts gives us :

$$\int_{\mathbb{R}^{2d}} u_i(s,y) D_{x_2}\tilde{q}(t,x;s,y)dy = -\int_{\mathbb{R}^{2d}} D_{y_2}u_i(s,y)\tilde{q}(t,x;s,y)dy,$$

so that :

$$\begin{aligned} D_{x_2}\tilde{u}_i(t,x) - D_{x_2}\tilde{u}_i(s,x) &= D_{x_2}\int_t^s \mathbb{E}\left[\phi_i(r, \tilde{X}_r^{t,x}) - (\mathcal{L} - \tilde{\mathcal{L}}^{t,\xi})\tilde{u}_i(r, \tilde{X}_r^{t,x})\right] dr \\ &\quad - \int_{\mathbb{R}^{2d}} \left(D_{x_2}u_i(s,y) - D_{x_2}u_i(s,x)\right)\tilde{q}(t,x;s,y)dy, \end{aligned}$$

since $D_{x_2}\tilde{u}_i(s,\cdot) = D_{x_2}u_i(s,\cdot)$ by the definition of $\tilde{u}_i$. On the one hand, we know from the proof of Proposition 1.1.3 in Chapter 1 that there exist a positive constant $C$ and a positive real $\gamma_{A.2.1}$, depending only on known parameters in **(H)** , such that

$$\left| D_{x_2}\int_t^s \mathbb{E}\left[\phi_i(r, \tilde{X}_r^{t,x}) - (\mathcal{L} - \tilde{\mathcal{L}}^{t,\xi})\tilde{u}_i(r, \tilde{X}_r^{t,x})\right] dr \right| \le C|s-t|^{\gamma_{A.2.1}}. \tag{A.3.6}$$

On the other hand we have that :

$$\begin{aligned}
&\int_{\mathbb{R}^{2d}} |D_{x_2}u_i(s,y) - D_{x_2}u_i(s,x)|\tilde{q}(t,x;s,y)dy \\
&\quad \leq \int_{\mathbb{R}^{2d}} \left(||D_{x_1x_2}u_i||_\infty |y_1 - x_1| + ||D_{x_2}u_i||_{.,.,\gamma_{A.2.1}} |y_2 - x_2|^{\gamma_{A.2.1}}\right) \hat{q}_c(t,x;s,y)dy \\
&\quad \leq C\int_{\mathbb{R}^{2d}} \left(|\Delta^2(\theta_{t,s}(x))y_2|^{\gamma_{A.2.1}} + |\Delta^1(\theta_{t,s}(x))y_1|\right) \hat{q}_c(t,x;s,y)dy \\
&\qquad +C\int_{\mathbb{R}^{2d}} \left(|\theta^2_{t,s}(x) - x_2|^{\gamma_{A.2.1}} + |\theta^1_{t,s}(x) - x_1|\right) \hat{q}_c(t,x;s,y)dy,
\end{aligned}$$

where $\Delta$ is the perturbation operator defined in 1.3.2. Then, we can find a positive constant $K(||F||_{\infty,\mathcal{K}})$, depending on the supremum norm of $F$ on $\mathcal{K}$ such that $|\theta^j_{t,s}(x) - x_j| \leq K(||F||_{\infty,\mathcal{K}})(s-t)$. We deduce that there exists a positive constant $K'(||F||_{\infty,\mathcal{K}})$ depending only on known parameters in **(H)** and $||F||_{\infty,\mathcal{K}}$ such that

$$\int_{\mathbb{R}^{2d}} |D_{x_2}u_i(s,y) - D_{x_2}u_i(s,x)|\tilde{q}(t,x;s,y)dy \leq K'(||F||_{\infty,\mathcal{K}})|s-t|^{\gamma_{A.2.1}}.$$

Together with (A.3.6), this concludes the proof for the first term in (A.3.3).

$\bullet$*Bound of* $||D^2_{x_1}u||_{\gamma_{A.2.1},.,.}$. As above, we write $\tilde{u} = (\tilde{u}_1, \tilde{u}_2)$ the solution of the PDE (A.2.1) with the boundary condition $u(s,.)$ at time $s \in (t,T)$. Let $i \in \{1,2\}$, in this case $\tilde{u}_i$ writes :

$$\tilde{u}_i(t,x) = \int_t^s \mathbb{E}\left[\phi_i(r, \tilde{X}^{t,x}_r) - (\mathcal{L} - \tilde{\mathcal{L}}^{t,\xi})\tilde{u}_i(r, \tilde{X}^{t,x}_r)\right]dr + \int_{\mathbb{R}^{2d}} u_i(s,y)\tilde{q}(t,x;s,y)dy.$$

Then,

$$\begin{aligned}
D^2_{x_1}\tilde{u}_i(t,x) &= D^2_{x_1}\int_t^s \mathbb{E}\left[\phi_i(r, \tilde{X}^{t,x}_r) - (\mathcal{L} - \tilde{\mathcal{L}}^{t,\xi})\tilde{u}_i(r, \tilde{X}^{t,x}_r)\right]dr \\
&\quad + \int_{\mathbb{R}^{2d}} u_i(s,y)D^2_{x_1}\tilde{q}(t,x;s,y)dy.
\end{aligned}$$

In this case, the proof of Proposition 1.3.1 shows that

$$D_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2) \neq -D_{y_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2),$$

and this does not allow to use integration by parts. This asymmetry comes from the transport of the initial condition of the first component in the second one. In order to recover the symmetry we write :

$$\begin{aligned}
\int_{\mathbb{R}^{2d}} u_i(s,y)D^2_{x_1}\tilde{q}(t,x;s,y)dy &= \int_{\mathbb{R}^{2d}} \left(\Delta^2(\theta_{t,s}(x))u_i\right)(s,y)D^2_{x_1}\tilde{q}(t,x_1,x_2;s,y_1,y_2)dy \\
&\quad + \int_{\mathbb{R}^{2d}} u_i(s,y_1,\theta^2_{t,s}(\xi))D^2_{x_1}\tilde{q}(t,x;s,y)dy \\
&= \int_{\mathbb{R}^{2d}} \left(\Delta^2(\theta_{t,s}(x))u_i\right)(s,y)D^2_{x_1}\tilde{q}(t,x;s,y)dy \\
&\quad + \int_{\mathbb{R}^{d}} u_i(s,y_1,\theta^2_{t,s}(\xi))D^2_{x_1}\tilde{q}_1(t,x_1;s,y_1)dy_1,
\end{aligned}$$

where $\tilde{q}_1$ is the marginal law of the first component in (A.1.1). This law is symmetric and Gaussian, so that :

$$\begin{aligned}\int_{\mathbb{R}^d} u_i(s, y_1, \theta^2_{t,s}(\xi)) D^2_{x_1} \tilde{q}_1(t, x_1; s, y_1) dy_1 &= \int_{\mathbb{R}^d} u_i(s, y_1, \theta^2_{t,s}(\xi)) D^2_{y_1} \tilde{q}_1(t, x_1; s, y_1) dy_1 \\ &= \int_{\mathbb{R}^d} D^2_{y_1} u_i(s, y_1, \theta^2_{t,s}(\xi)) \tilde{q}_1(t, x_1; s, y_1) dy_1.\end{aligned}$$

By following the arguments of the proof of **(iii)** we can deduce that for all $(t, x)$ in $\mathcal{K}$, the mapping $t \mapsto D^2_{x_1} u(t, x)$ is Hölder continuous. This concludes the proof of the second term in (A.3.3).

• *Bound of* $||D_{x_1} u||_{\gamma_{A.2.1},\cdot,\cdot}$. Same arguments lead to the Hölder continuity in time for the first order derivative $D_{x_1} u$.

□

# CHAPITRE 2

## A cubature based algorithm to solve decoupled McKean-Vlasov Forward Backward Stochastic Differential Equations

We call decoupled McKean-Vlasov forward backward stochastic differential equation (MKV-FBSDE) the following FBSDE system :

$$\begin{cases} dX_t^x = \sum_{i=0}^d V_i(t, X_t^x, \mathbb{E}\varphi_i(X_t^x))dB_t^i \\ dY_t^x = -f(t, X_t^x, Y_t^x, Z_t^x, \mathbb{E}\varphi_f(X_t^x, Y_t^x))dt + Z_t^x dB_t^{1:d} \\ X_0^x = x, \quad Y_T^x = \phi(X_T^x) \end{cases} \tag{2.0.7}$$

for any $t$ in $[0,T]$, $T > 0$ be given. We place ourselves in a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$, with $B_t^{1:d}$ a $d$-dimensional adapted Brownian motion and $B_t^0 = t$. We take $V_i : (t, y, w) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} \mapsto V_i(t, y, w)$ ; functions $\varphi_i : y \in \mathbb{R}^d \mapsto \varphi_i(y) \in \mathbb{R}$, $i = 0, \cdots, d$ and $\varphi_f : (y, y') \in \mathbb{R}^d \times \mathbb{R} \mapsto \varphi_f(y, y')$ and the mapping $f : t, y, y', z, w \in [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \mapsto f(t, y, y', z, w) \in \mathbb{R}$ to be bounded and infinitely differentiable with bounded derivatives. The mapping $\phi$ is an at least Lipschitz function from $\mathbb{R}^d$ to $\mathbb{R}$ whose precise regularity is given below.

McKean Vlasov processes may be regarded as a limit approximation for interacting systems with large number of particles. They appeared initially in statistical mechanics, but are now used in many fields because of the wide range of applications requiring large populations interactions. For example, they are used in finance, as factor stochastic volatility models [Ber09] or uncertain volatility models [GHL11] ; in economics, in the theory of "mean field games" recently developed by J.M. Lasry and P.L. Lions in a series of papers [LL06b, LL06a, LL07] (see also [CDL12, CD12b, CD12a] for the probabilistic counterpart) and also in physics, neuroscience, biology, etc. In section 2.4, we present a class of control problems in which equation (2.0.7) explicitly appears.

The note of Sznitman [Szn91] gives a complete overview on the topic of systems with a large number of particles. A proof of the existence and uniqueness of the solution of a MKV-FBSDE system related but different to the one of our setup is found in [BLP09]. These existence and uniqueness results are easily extended to (2.0.7).

**A cubature algorithm for MKV-FBSDE processes.** Cubature on Wiener space was introduced in 2004 by T.Lyons and N.Victoir [LV04], following the earlier work of S.

Kusuoka [Kus01]. Ever since, the cubature method has been used to solve the problem of calculating Greeks in finance [Tei06], non-linear filtering problems [CG07], stochastic partial differential equations [BT08], [DTV12] and stochastic backward differential equations [CM10a, CM10b].
The main idea of the cubature method consists in replacing the Brownian motion by choosing randomly a path among an a priori (finite) set [32] of continuous functions from $[0, T]$ to $\mathbb{R}^d$ with bounded variations such that the expectation of the iterated integrals against both the Brownian and such paths are the same, up to a given order $m$. Hence, the SDE is replaced by a system of weighted ODEs.

We give the main idea to construct a cubature based approximation scheme for (2.0.7). The main issue in the case of a MKV-FBSDE is the McKean-Vlasov dependence that appears in the coefficients. This dependence breaks the Markov property (considered only on $\mathbb{R}^d$) of the process so that it is not possible to apply, a priori, many classical analysis tools. In order to handle this problem, the idea consists in taking benefit on the following observation : *given the law of the solution of the system,* (2.0.7) *is a classical time inhomogeneous FBSDE* (the law just acts as a time dependent parameter).

Let $(\eta_t)_{0 \leq t \leq T}$ be a family of probability measures on $\mathbb{R}^d$, and let us fix the law in the McKean-Vlasov terms of (2.0.7) to be $(\eta_t)_{0 \leq t \leq T}$ . For this modified system, we may apply a classical cubature FBSDE scheme for the forward component (the time dependence of the coefficients being handled as an additional dimension). The trick consists in taking advantage of the decoupled setting : we first build a cubature tree (depending on the order of the cubature) and then go back along the nodes of the tree by computing the current value of the backward process as a conditional expectation at each node. We refer to [CM10b] or [CM10a] for a detailed description of such algorithm.

Obviously, at each step of the scheme, we pay the price of using an arbitrary probability measure as parameter for the coefficients instead of the law of the process. Therefore this law has to be chosen carefully in order to keep a good control on the error and achieve convergence. An example of a "good choice" is to take at each step of the cubature tree the discrete marginal law given by the solution of the ODEs along the cubature paths and corresponding weights. We show that for a cubature of order $m$ and a number $N$ of discretization steps, this choice of approximation law leads to a $N^{-(m-1)/2}$ order approximation of the expectation of any bounded and $m + 2$ times continuously differentiable functional of the forward component, when all the derivatives are bounded [33], and to a first order approximation scheme of the backward component, where the given orders stand for the supremum norm error. Higher orders of approximation are also obtained by correcting some terms in the algorithm.

As it is pointed out in [LV04] and [CM10a], the regularity of the terminal condition $\phi$ in (2.0.7) may be relaxed to Lipschitz and the approximation convergence rate preserved, provided that the vector fields are uniformly non-degenerate (in fact, the condition in the

32. Explicit examples of such functions are given in [LV04]. We put back this discussion to a next subsection.

33. this is the special case when $\phi$ is $m + 2$ times continuously differentiable with bounded derivatives and $f = 0$ in (2.0.7).

given references is weaker, since the vector fields are supposed to satisfy an UFG condition, see [KS87]). This relies on the regularization properties of parabolic and semi-linear parabolic PDEs (see [Fri08] for an overview in the elliptic case and respectively [KS87] and [CD12c] for the UFG case). We show that this remains true in the McKean-Vlasov case and that the convergence rate still holds when the function $\phi$ is Lipschitz only and when the vector fields are uniformly elliptic.

Usually, forward MKV-SDEs are solved by using particle algorithms (see for example [AKH02, TV03] or [Bos05] and references therein) in which the McKean term is approached with the empirical measure of a large number of interacting particles with independent noise. Adapting such algorithms to the forward-backward problem is not obvious as the high dimension of the involved Brownian motion (given by the number of particles) induces, a priori, a high dimension backward problem with the obvious consequences for the numerical implementation. In comparison, our proposed algorithm gives a deterministic approximation of the McKean term, and since it does not induce any additional noise, it does not increase the dimension of the backward problem.

Although our algorithm works for decoupled MKV-FBSDEs, we believe this solver may also be considered as a building block if one is interested in approaching the fully coupled case (when the forward coefficients depend on the backward variable), for example via fixed point procedures. Nevertheless, a lot of work is required to define the precise conditions and setup in which such algorithm would converge to the (or at least a) solution of the fully coupled problem.

**The conditional system.** Let us shortly develop what we mean with the sentence "*given the law of the solution of the system,* (2.0.7) *is a classical time inhomogeneous FBSDE*". Working with a non-linear problem, such as MKV-FBSDE, could be tricky. In our case, the main object to work with is the *conditional system*. This is the formulation that allows to get rid of the dependence on the law and to replace it by a time dependent parameter.

Following the same line of arguments presented in Buckdahn et al. [BLP09], it is well seen that there exists a unique solution $\{X_t^x, Y_t^x\}_{t\geq 0}$ to the system (2.0.7). In a Markovian setting, the law of this couple is entirely determined by the law $\mu = (\mu_t)_{0\leq t\leq T}$ of the forward process $(X_t)_{0\leq t\leq T}$ and a given deterministic function $u : [0,T]\times\mathbb{R}^d \to \mathbb{R}$. In our case, one can show that this remains true (see Section 2.7 below for a proof) so that there exists a deterministic $u(t,y)$ such that for all $t$,

$$Y_t = u(t, X_t). \tag{2.0.8}$$

We prove that under appropriate assumptions $u$ is regular and satisfies the parametrized non-local semi linear PDE :

$$\begin{cases} D_t u(t,y) + \mathcal{L}^\mu u(t,y) = f\left(t, y, u(t,y), (\mathcal{V}^\mu u(t,y))^T, \langle \mu_t, \varphi_f(\cdot, u(t,\cdot))\rangle\right) \\ u(T,y) = \phi(y) \end{cases}, \tag{2.0.9}$$

where $\mathcal{V}^\mu u$ stands for the row vector $(\nabla u \cdot V_1, \ldots, \nabla u \cdot V_d)$, $(\mathcal{V}^\mu u)^T$ is the transpose of $\mathcal{V}^\mu u$ and $\mathcal{L}^\mu$ is the generator of the forward component in (2.0.11) below and given by :

$$\mathcal{L}^\mu := V_0(\cdot,\cdot,\langle \mu_\cdot, \varphi_0\rangle)\cdot D_y + \frac{1}{2}\mathrm{Tr}[VV^T(\cdot,\cdot,\langle \mu_\cdot, \varphi_i\rangle)D_y^2]. \tag{2.0.10}$$

Here, we used the duality notation $\langle \mu, \varphi_i\rangle$ for $\int \varphi_i d\mu$. Likewise, the superscript $\mu$ means that the vector fields are taken at the point $(\cdot,\cdot,\langle \mu_\cdot, \varphi_i\rangle)$ (where the $i\in\{0,\cdots,d\}$ is taken with respect to the corresponding vector field), $V$ is the matrix $[V_1,\cdots,V_d]$, "$\cdot$" stands for the euclidean scalar product on $\mathbb{R}^d$ and "Tr" for the trace.

The conditional MKV-FBSDE system is then defined as

$$\begin{cases} dX_s^{t,y,\mu} = \sum_{i=0}^d V_i(s, X_s^{t,y,\mu}, \langle \mu_s, \varphi_i \rangle) dB_s^i \\ dY_s^{t,y,\mu} = -f(s, X_s^{t,y,\mu}, Y_s^{t,y,\mu}, Z_s^{t,y,\mu}, \langle \mu_s, \varphi_f(\cdot, u(s,\cdot))\rangle) ds + Z_s^{t,y} dB_s^{1:d} \\ X_t^{t,y,\mu} = y, \quad Y_T^{t,y,\mu} = \phi(X_T^{t,y,\mu}). \end{cases} \tag{2.0.11}$$

Let us remark that in this setting, we can also define the deterministic mapping $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ such that

$$Z_t = v(t, X_t). \tag{2.0.12}$$

As in the classical BSDE theory, under appropriate regularity conditions,

$$v(t,x) = (\mathcal{V}^\mu u(t,x))^T.$$

**Assumptions**. As the reader might guess from the previous discussion, the error analysis of the proposed algorithm uses extensively the regularity of $(t,x) \in [0,T] \times \mathbb{R}^d \mapsto \mathbb{E}\phi(X_T^{t,x,\mu})$ for the forward part and of the solution $u$ of (2.0.9) for the backward part. Therefore, we present two types of hypotheses that guarantee that the required regularity is attained.

The first option we present is to require smoothness on the boundary condition and all the coefficient functions, from where we will deduce the necessary regularity. However, it is also interesting to consider boundary conditions with less regularity. In this case, we need to compensate the regularity loss by imposing stronger diffusion conditions on the forward variable, namely asking for uniform ellipticity of the diffusion matrix $V$.

**(SB):** We say that assumption **(SB)** holds if the mapping $\phi$ in (2.0.7) is $C_b^\infty$.

**(LB):** We say that assumption **(LB)** holds if the mapping $\phi$ in (2.0.7) is uniformly Lipschitz continuous and if the matrix $VV^T$ is uniformly elliptic i.e., there exists $c > 0$ such that

$$\forall (t,y,w) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}, \forall \varsigma \in \mathbb{R}^d, c^{-1}|\varsigma|^2 \le VV^T(t,y,w)\varsigma \cdot \varsigma \le c|\varsigma|^2.$$

**Remark.** *A reader familiarized with the cubature method might wonder why we assume uniform ellipticity instead of the weaker UFG condition usually needed for applying the method with Lipschitz boundary conditions. The reason is that the smoothing results of Kusuoka and Stroock [KS87] hold for space dependent vector fields only, and therefore do not apply directly to our framework with a time dependence coming from the McKean term. There is some extension in the time inhomogeneous case that do not include derivatives in the $V_0$ direction (see for example [CM02] and references therein), but, to the best of our knowledge, there is no result that could be applied to our framework.*

**Towards a more general class of coefficients.** We have chosen to work with the assumed explicit dependence of the coefficients with respect to the law, as it is very natural in practice. In fact as a consequence of our analysis, our algorithm works for a more general class of MKV-FBSDE, i.e. for a system written as

$$\begin{cases} dX_t^x = \sum_{i=0}^d V_i(t, X_t^x, \mu_t) dB_t^i \\ dY_t^x = -f(t, X_t^x, Y_t^x, Z_t^x, \mu_t^{X,Y}) dt + Z_t^x dB_t^{1:d} \\ X_0^x = x, \quad Y_T^x = \phi(X_T^x), \end{cases} \tag{2.0.13}$$

where $\mu^{X,Y} = (\mu_t^{X,Y})_{0 \le t \le T}$ denotes the joint law of $(X_t, Y_t)_{0 \le t \le T}$, the coefficients $V_i$, $0 \le i \le d$ (and $f$) are Lipschitz continuous with respect to an appropriately defined distance in the space of probability measures on $\mathbb{R}^d$ (respectively $\mathbb{R}^d \times \mathbb{R}$ ) which integrates the square of the usual Euclidean norm. The distance we consider is defined by duality :

let $\mathcal{F}$ be a (sufficiently rich [34]) class of functions (that will be detailed in the following). Then we define the distance $d_{\mathcal{F}}$ between two probability measures on $\mathbb{R}^n$ by

$$d_{\mathcal{F}}(\mu, \nu) = \sup_{\varphi \in \mathcal{F}} |\langle \varphi, \mu - \nu \rangle|. \tag{2.0.14}$$

In this decoupled case the Lipschitz property of the coefficients with respect to $d_{\mathcal{F}}$ ensures the existence of a unique solution of (2.0.13) [35]. Then, we are able to analyze the convergence of our procedure in two different cases. When $\mathcal{F}$ is the class of 1-Lipschitz functions, i.e. when the distance is the so-called Wasserstein-1 distance, and when the vector fields are uniformly elliptic our algorithm leads to an $N^{-1/2}$ order approximation [36]. When $\mathcal{F}$ is the class of $C_b^\infty$ functions we obtain an $N^{-1}$ order approximation without any ellipticity assumption on the diffusion matrix.

**Objectives and organization of this Chapter.** As a corollary of the discussion on the conditional system we can resume our objective as the approximation of $\mathbb{E}\phi(X_T^x)$, where $(X_t^x)_{0 \leq t \leq T}$ is the solution of (2.0.7) and of $u$ satisfying (2.0.8).

This Chapter is organized as follows : Section 2.1 states the algorithm, while the convergence rate of the forward and backward approximations is stated in Section 2.2. Then, we give a numerical example for each set of hypotheses **(SB)** and **(LB)** in Section 2.3. A class of control problems is introduced in Section 2.4. The remainder of the Chapter is dedicated to the proof of the convergence. For the sake of simplicity, we first recall some definitions, basic facts and notations in Section 2.5. The forward and backward convergence rates for regular boundary conditions are successively proved in Section 2.6 and the common mathematical tools are given in Section 2.7. Section 2.8 presents the extension to the Lipschitz boundary condition case and the announced generalization of the law dependence of the McKean terms.

**Notations.** As we are treating with objects exhibiting different dependences, the notation can become a bit heavy. For the sake of simplicity, we adopt the following conventions. We denote by $\varphi$ the function $\varphi = [\varphi_1, \cdots, \varphi_d]$. For two positive integers $i < j$, the notation "$i : j$" means "from index $i$ to $j$". For all $\varsigma \in \mathbb{R}^n$, $n \in \mathbb{N}$ the partial derivative $[\partial / \partial \varsigma]$ is denoted by $\partial_\varsigma$. Let $g : \ y \in \mathbb{R}^d \mapsto g(y) \in \mathbb{R}$ be a $p$-continuously differentiable function. We set $||g||_{\infty,p} := \max_{j \leq p} ||\partial_y^j g||_\infty$. We say that a function $g$ from $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ is $C_b^p$, $p \in \mathbb{N}^*$ if it is bounded and $p$-times continuously differentiable with bounded derivatives. We usually denote by $\eta$ (eventually with an exponent) a family of probability measures $(\eta_t)_{0 \leq t \leq T}$ on $\mathbb{R}^d$. For such a family, we set $\mathcal{L}^\eta$ to be the second order operator of the form (2.0.10) with $\eta$ instead of $\mu$. In general, we will work with the vector fields taken at the point $(\cdot, \cdot, \langle \mu_\cdot, \varphi_i \rangle)$ (where the $i \in \{0, \cdots, d\}$ signals the corresponding vector field) and in general we will omit the explicit dependence on $\mu$ in the notation. In any case, we will mark the law dependence explicitly when needed, in particular when a dependence with respect to a different law appears.

34. It should contain the space of 1-Lipschitz functions.
35. by definition of $\mathcal{F}$, the distance $d_{\mathcal{F}}$ is less than or equal to the Wasserstein 2 distance. Then, one uses the same kind of arguments as in [Szn91].
36. Recall that $N$ denotes the number of discretization steps.

### *2.1* Algorithm

**Multi-index.** Multi-indices allow to easily manage differentiation and integration in several dimensions. Let

$$\mathcal{M} = \{\emptyset\} \cup \bigcup_{l \in \mathbb{N}^*} \{0, 1, \ldots, d\}^l, \tag{2.1.1}$$

denotes the set of multi-indices where $\emptyset$ refers, for the sake of completeness, to the zero-length multi-index. We define "$*$" to be the concatenation operator such that if $\beta^1 = (\beta^1_1, \ldots, \beta^1_l)$ and $\beta^2 = (\beta^2_1, \ldots, \beta^2_n)$ then $\beta^1 * \beta^2 = (\beta^1_1, \ldots, \beta^1_l, \beta^2_1, \ldots, \beta^2_n)$.

**Cubature on Wiener Space.** In the introduction, we mentioned that the cubature method consists in replacing the Brownian path by choosing randomly a path $\omega$ among a finite subset $\{\omega_1, \cdots, \omega_\kappa\}$, $\kappa \in \mathbb{N}^*$, of $C^0_{\mathrm{bv}}([0,T], \mathbb{R}^d)$ (the set of continuous functions from $[0,T]$ to $\mathbb{R}^d$ with bounded variations) with probability $\lambda$ in $\{\lambda_1, \cdots, \lambda_\kappa\} \subset \mathbb{R}^+$. We precise this notion with the definition given by Lyons and Victoir [LV04] :

**Definition 2.1.1.** *Let $m$ be a natural number and $t \in \mathbb{R}^+$. A $m$-cubature formula on the Wiener space $C^0([0,t], \mathbb{R}^d)$ is a discrete probability measure $\mathbb{Q}_t$ with finite support on $C^0_{\mathrm{bv}}([0,t], \mathbb{R}^d)$ such that the expectation of the iterated Stratonovitch integrals of degree $m$ under the Wiener measure and under the cubature measure $\mathbb{Q}_t$ are the same, i.e., for all multi-index $(i_1, \cdots, i_l) \in \{1, \cdots, d\}^l$, $l \le m$*

$$\begin{aligned} \mathbb{E} \int_{0<t_1<\cdots<t_l<t} \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_l}_{t_l} &= \mathbb{E}_{\mathbb{Q}_t} \int_{0<t_1<\cdots<t_l<t} \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_l}_{t_l} \\ &= \sum_{j=1}^{l} \lambda_j \int_{0<t_1<\cdots<t_l<t} d\omega^{i_1}_j(t_1) \cdots d\omega^{i_l}_j(t_l), \end{aligned}$$

*where "$\circ$" stands for the Stratonovich operator and $\omega^i_j$ for the $i^{\text{th}}$ coordinate of the $j^{\text{th}}$ path.*

As a direct consequence of the Taylor-Stratonovitch expansion, a cubature formula of degree $m$ is such that :

$$|(\mathbb{E} - \mathbb{E}_{\mathbb{Q}_t}) F(B^{1:d}_t)| \le C t^{(m+1)/2} ||F||_{m+2,\infty}, \tag{2.1.2}$$

for all bounded and $m+2$ times continuously differentiable function $F$ with bounded derivatives.

Of course this error control is not in general small, but the Markovian and scaling properties of the Brownian motion can be used to apply the cubature method iteratively in small subdivisions of the interval $[0,t]$ for which we have a good error control.

Indeed, consider a cubature formula $\mathbb{Q}_1$ of order $m \in \mathbb{N}^*$ with support $\{\omega_1, \ldots, \omega_\kappa\}$ and corresponding weights $\{\lambda_1, \ldots, \lambda_\kappa\}$. For all $h > 0$ and any $t \in [0, T-h]$, one can deduce a cubature measure $\mathbb{Q}_{t,t+h}$ of order $m$ with finite support on $C^0_{\mathrm{bv}}([t, t+h], \mathbb{R}^d)$ equal to $\{\tilde{\omega}_1, \cdots, \tilde{\omega}_\kappa\}$ with the same weights $\{\lambda_1, \ldots, \lambda_\kappa\}$ and where the paths are defined as $\tilde{\omega}_j : s \in [t, t+h] \mapsto \tilde{\omega}(s) = \sqrt{h}\omega_j((s-t)/h)$ for all $1 \le j \le \kappa$.

Then, by virtue of the Markovian property, this subdivision leads to the construction of a tree which has $\kappa^k$ nodes (corresponding to the number of paths) at the $k^{\text{th}}$ subdivision. Each path $\tilde{\omega}_{(i_1,\ldots,i_k)}$, where $(i_1, \ldots, i_k)$ stands for the trajectory of the path, has then a cumulate weight of the form $\Lambda_{(i_1,\ldots,i_k)} = \prod_{j=1}^k \lambda_{i_j}$, see the example and figure 1 below.

### *2.1.1* Main idea

Take a subdivision of the time interval $[0, T]$ into $N \in \mathbb{N}^*$ steps $0 = T_0 < \cdots < T_N = T$. The procedure can be decomposed in two parts :

(1) **Building the tree $\mathcal{T}$.** This part of our algorithm can be resumed as a combined Euler-cubature approach and can be divided in four steps.

   (a) First, freeze in the space of probability measures the law that appears in the coefficients of (2.0.7) (the choice of this measure is explained below). At step 0, this measure is the Dirac mass in the starting point.

   (b) Freeze, in time, the given deterministic measure : this is an Euler step.

   (c) Apply the cubature method. This will produce a cloud of deterministic particles given by the solution of the resulting ODEs

   (d) At each step, construct a discrete measure coming from the obtained cloud of particles and their associated cumulative weights. This is the law to be used to approximate the law in the coefficients of (2.0.7).

   The reader might guess that the order of approximation of such an algorithm is one, due to the Euler step. Hence, in order to obtain higher order, we expand the function that appears in the McKean-Vlasov part up to a certain order, which is denoted by "$q$" in the sequel.

(2) **The backward component**. The backward component of the algorithm runs by assigning the value of the function on the boundary (known from the definition of the equation), and then back-propagating its value thanks to

   (a) A discretization scheme for the backward approximation

   (b) The cubature measure for finding conditional expectations.

   As mentioned before, we present two versions of the algorithm, with convergence of order one and two. As will be clear in the definition, the change in convergence order requires the use of a different backward scheme and cubature order.

**Example : one dimensional cubature of order $m = 3$.** In this case, we may use a cubature formula with $\kappa = 2$ paths given by $\{+t, -t\}$, and associated weights : $\{\lambda_1 = 1/2, \lambda_2 = 1/2\}$. Let us explain the idea behind the algorithm with an example for 2 steps, as shown in Figure 1. We initialize the tree at a given point $x$, and the law at $T_0$ as $\delta_x$. Then, we find two descendants given as the solution of an ODE that uses the position $X$, the two cubature paths, and an approximated law using the information at time 0. Each descendant will have a weight equal to the product of the weight of its parent times the weight given to the corresponding cubature path. Once all nodes at time $T_1$ are calculated, we obtain the discrete measure $\hat{\mu}_{T_1}$, the law approximation at time $T_1$. The process is then repeated for each node at time $T_1$ to reach the final time $T = T_2$.

Figure 1 right illustrates the idea behind the backward approximation : the approximated function $\hat{u}$ is defined first at the leaves of the constructed tree, and then back-propagates using the approximated law to obtain $\hat{u}$ at previous times. The back-propagation is made by conditional expectation : average with respect to the weight of each cubature path.

### *2.1.2* Algorithms

Having the general idea in mind, we can give a precise description of each of the two main parts of our proposed algorithm. Since the cubature involves Stratanovitch integrals,

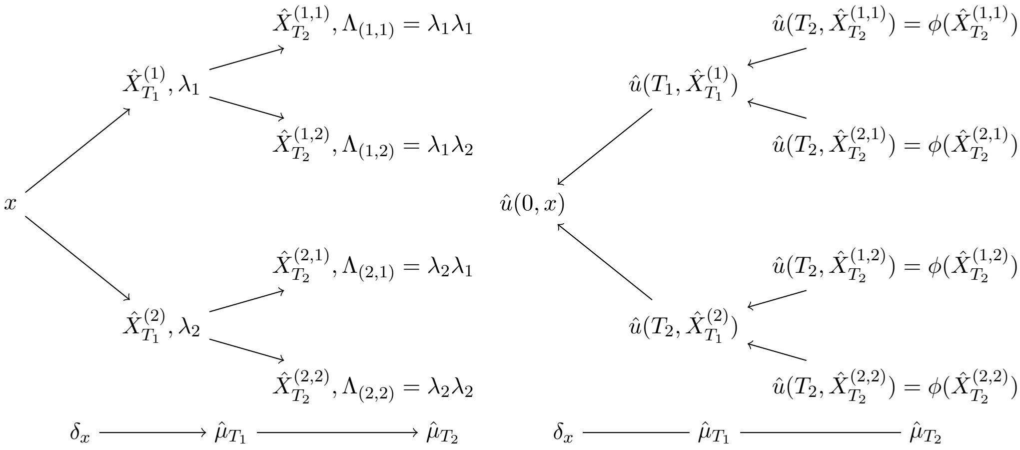


FIGURE 1. Left : Cubature tree. Right : Backward scheme.

we set :

$$\bar{V}_0^k = V_0^k - \frac{1}{2}\sum_{i,j=1}^{d} V_j^i \frac{\partial}{\partial x_j} V_k^j, \tag{2.1.3}$$

for all $k \in \{1, \cdots, d\}$ and rewrite the system (2.0.7) as :

$$\begin{cases} dX_t^x = \bar{V}_0(t, X_t^x, \mathbb{E}\varphi_0(X_t^x))dt + \sum_{i=1}^d V_i(t, X_t^x, \mathbb{E}\varphi_i(X_t^x)) \circ dB_t^i \\ dY_t^x = -f(t, X_t^x, Y_t^x, Z_t^x, \mathbb{E}\varphi_f(X_t^x, Y_t^x))dt + Z_t^x dB_t^{1:d} \\ X_0^x = x, \quad Y_T^x = \phi(X_T^x). \end{cases} \tag{2.1.4}$$

In order to make the description of the algorithm as clear as possible, for any $k, \kappa$ in $\mathbb{N}$, we set $\mathcal{S}_\kappa(k) = \{\text{multi-index } (j_1, \cdots, j_k) \in \{1, \cdots, \kappa\}^k\}$, i.e., $\mathcal{S}_\kappa(k)$ is the set of multi-indices with entries between $1, \ldots, \kappa$ of length (exactly) $k$.

***2.1.2.1 Building the tree $\mathcal{T}(\gamma, q, m)$*** **The subdivision.** Let $\gamma > 0$, $N \in \mathbb{N}^*$, let $0 = T_0 < \ldots < T_N = T$ be a discretization of the time interval $[0, T]$ given as

$$T_k = T\left[1 - \left(1 - \frac{k}{N}\right)^\gamma\right] \tag{2.1.5}$$

and let $\Delta_{T_k} = T_k - T_{k-1}$.

**Remark.** *When the boundary condition is not smooth, we take a non-uniform subdivision in order to refine the discretization step close to the boundary as proposed by Kusuoka in [Kus01]. If, on the contrary, the boundary condition is smooth, we may use a classical uniform discretization. For this reason, in the following we assume that $\gamma = 1$ if **(SB)** holds, and that $\gamma > m - 1$ under **(LB)**.*

Let $\gamma$ be given as explained above, $q$ and $m$ be two given integers, and $\big\{\{\omega_1, \cdots, \omega_\kappa\}, \{\lambda_1, \cdots, \lambda_\kappa\}\big\}$ be a $m$ order cubature (the number $\kappa$ of paths and weights depends on $m$). Recall that $\omega_j : t \in [0, 1] \mapsto (\omega_j^1(t), \ldots, \omega_j^d(t)) \in \mathbb{R}^d$ is some continuous function with bounded variation and for all $t$ in $[0, T]$, we set $\omega_0(t) = t$. Examples of cubature formulas of order $3, 5, 7, 9, 11$ can be found in [LV04] or [GL11].

**Algorithm 1** Cubature Tree $\mathcal{T}(\gamma, q, m)$

Set $(X^{\emptyset}, \hat{\mu}_{T_0}, \Lambda_0) = (x, \delta_x, 1)$
**for** $0 \leq i \leq d$ **do**
  Set $F_i(t, \hat{\mu}_{T_0}) = \sum_{p=0}^{q-1} \frac{1}{p!}(t - T_0)^p \langle \delta_x, (\mathcal{L}^{\delta_x})^p \varphi_i \rangle$
**end for**
**for** $1 \leq k \leq N - 1$ **do**
  **for** $\pi \in \mathcal{S}_\kappa(k)$ **do**
    **for** $1 \leq j \leq \kappa$ **do**
      Define $\hat{X}^{\pi * j}_{T_{k+1}}$ as the solution of the ODE :

$$d\hat{X}^{\pi*j}_t = \sum_{i=0}^{d} V_i(t, \hat{X}^{\pi*j}_t, F_i(t, \hat{\mu}_{T_k}))\sqrt{\Delta_{T_{k+1}}} d\omega^i_j((t - T_k)/\Delta_{T_{k+1}}),$$
$$\hat{X}^{\pi*j}_{T_k} = \hat{X}^{\pi}_{T_k}$$

      Set the associated weight : $\Lambda_{\pi * j} = \Lambda_\pi \lambda_j$
    **end for**
  **end for**
  Set $\hat{\mu}_{T_{k+1}} = \sum_{\pi \in \mathcal{S}_\kappa(k+1)} \Lambda_\pi \delta_{\hat{X}^{\pi}_{T_{k+1}}}$
  **for** $0 \leq i \leq d$ **do**
    Set $F_i(t, \hat{\mu}_{T_{k+1}}) = \sum_{p=0}^{q-1} \frac{1}{p!}(t - T_k)^p \langle \hat{\mu}_{T_{k+1}}, (\mathcal{L}^{\hat{\mu}})^p \varphi_i \rangle$
  **end for**
**end for**

#### *2.1.2.2 Backward scheme*

**Algorithm 2** First order backward scheme

**for** $\pi \in \mathcal{S}_\kappa(N)$ **do**
  Set $\hat{u}^1(T_N, \hat{X}^{\pi}_{T_N}) = \phi(\hat{X}^{\pi}_{T_N})$
  Set $\hat{v}^1(T_N, \hat{X}^{\pi}_{T_N}) = 0$
**end for**
**for** $N - 1 \geq k \geq 1$ **do**
  **for** $\pi \in \mathcal{S}_\kappa(k)$ **do**
    $\hat{v}^1(T_k, \hat{X}^{\pi}_{T_k}) = \frac{1}{\Delta_{T_{k+1}}} \sum_{j=1}^{\kappa} \lambda_j \hat{u}^1(T_{k+1}, X^{\pi*j}_{T_{k+1}}) \sqrt{\Delta_{T_{k+1}}} \omega_j(1)$
    **for** $1 \leq j \leq \kappa$ **do**
      $\hat{\Theta}^{\pi,1}_{k+1,k}(j) = \Big(T_{k+1}, \hat{X}^{\pi*j}_{T_{k+1}}, \hat{u}^1(T_{k+1}, \hat{X}^{\pi*j}_{T_{k+1}}), \hat{v}^1(T_k, \hat{X}^{\pi}_{T_k}), F^1(T_{k+1}, \hat{\mu}_{T_{k+1}})\Big)$
    **end for**
    $\hat{u}^1(T_k, \hat{X}^{\pi}_{T_k}) = \sum_{j=1}^{\kappa} \lambda_j \left( \hat{u}^1(T_{k+1}, X^{\pi*j}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\hat{\Theta}_{k+1,k}(j)) \right)$
  **end for**
  Set $F^1(T_{k+1}, \hat{\mu}_{T_{k+1}}) = \langle \hat{\mu}_{T_{k+1}}, \varphi_f(\cdot, \hat{u}^1(T_{k+1}, \cdot)) \rangle$
**end for**

**Algorithm 3** Second order backward scheme

**for** $\pi \in \mathcal{S}_\kappa(N)$ **do**
  Set $\hat{u}^2(T_N, \hat{X}^\pi_{T_N}) = \phi(\hat{X}^\pi_{T_N})$
  Set $\hat{v}^2(T_N, \hat{X}^\pi_{T_N}) = 0$
**end for**
**for** $\pi \in \mathcal{S}_\kappa(N-1)$ **do**
  Set $\hat{u}^2(T_{N-1}, \hat{X}^\pi_{T_{N-1}}) = \hat{u}^1(T_{N-1}, \hat{X}^\pi_{T_{N-1}})$ and $\hat{v}^2(T_{N-1}, \hat{X}^\pi_{T_{N-1}}) = \hat{v}^1(T_{N-1}, \hat{X}^\pi_{T_{N-1}})$
  Set $F^2(T_{N-1}, \hat{\mu}_{T_{N-1}}) = \langle \hat{\mu}_{T_{N-1}}, \varphi_f(\cdot, \hat{u}^2(T_{N-1}, \cdot))\rangle$
**end for**
**for** $N-2 \geq k \geq 1$ **do**
  **for** $\pi \in \mathcal{S}_\kappa(k)$ **do**
    **for** $1 \leq j \leq \kappa$ **do**
      $\hat{\Theta}^{\pi,2}_{k+1}(j) = \Big(T_{k+1}, \hat{X}^{\pi*j}_{T_{k+1}}, \hat{u}^2(T_{k+1}, \hat{X}^{\pi*j}_{T_{k+1}}), \hat{v}^2(T_{k+1}, \hat{X}^{\pi*j}_{T_{k+1}}), F^2(T_{k+1}, \hat{\mu}_{T_{k+1}})\Big)$
      $\hat{\zeta}^{\pi*j}_{k+1} := 4\dfrac{1}{\Delta_{T_{k+1}}}\sqrt{\Delta_{T_{k+1}}}\omega_j(1) - 6\dfrac{1}{\Delta^2_{T_{k+1}}}\displaystyle\int_{T_k}^{T_{k+1}}(s-T_k)\sqrt{\Delta_{T_{k+1}}}d\omega_j((s-T_k)/\Delta_{T_{k+1}})$
    **end for**
    Set $\hat{v}^2(T_k, \hat{X}^\pi_{T_k}) = \displaystyle\sum_{j=1}^{\kappa}\lambda_j\left(\hat{u}^2(T_{k+1}, X^{\pi*j}_{T_{k+1}}) + \Delta_{T_{k+1}}f(\hat{\Theta}^{\pi*j,2}_{k+1})\right)\hat{\zeta}^{\pi*j}_{k+1}$
    [*Predictor*] $\tilde{u}(T_k, \hat{X}^\pi_{T_k}) = \displaystyle\sum_{j=1}^{\kappa}\lambda_j\left(\hat{u}^2(T_{k+1}, X^{\pi*j}_{T_{k+1}}) + \Delta_{T_{k+1}}f(\hat{\Theta}^{\pi*j,2}_{k+1})\right)$
    Set $\tilde{F}(T_k, \hat{\mu}_{T_k}) = \langle \hat{\mu}_{T_k}, \varphi_f(\cdot, \tilde{u}(T_k, \cdot))\rangle$
    [*Corrector*] $\tilde{\Theta}^\pi_k = \Big(T_k, \hat{X}^\pi_{T_k}, \tilde{u}(T_k, \hat{X}^\pi_{T_k}), \hat{v}^2(T_k, \hat{X}^\pi_{T_k}), \tilde{F}(T_k, \hat{\mu}_{T_k})\Big)$
    $\hat{u}^2(T_k, \hat{X}^\pi_{T_k}) = \displaystyle\sum_{j=1}^{\kappa}\lambda_j\left(\hat{u}^2(T_{k+1}, X^{\pi*j}_{T_{k+1}}) + \frac{1}{2}\Delta_{T_{k+1}}\left(f(\hat{\Theta}^{\pi*j,2}_{k+1}) + f(\tilde{\Theta}^\pi_k)\right)\right)$
    Set $F^2(T_k, \hat{\mu}_{T_k}) = \langle \hat{\mu}_{T_k}, \varphi_f(\cdot, \hat{u}^2(T_k, \cdot))\rangle$
  **end for**
**end for**

**Remark.** *The initialization value for $v$ at the boundary, that we have fixed in $0$, is arbitrary, given that the first steps in both algorithms does not use this value.*

*However, if the algorithm is used under **(SB)** and the values of $D_x u$ can be easily calculated on the boundary, we have a natural initialization value for $v$. In this case, we may initialize the backward algorithms of order one and two to reflect this additional information, by setting $v(t,x) = (\mathcal{V}^\mu u(t,x))^T$.*

*This modification will have no effect at all for the first order scheme, and is interesting only from the point of view of consistence. On the other hand, on the second order algorithm, the natural initialization of $v$ allows to skip the first order step. This change does not affect the overall rate of convergence of the algorithm but will induce a reduction in the error constant whence of the total approximation error.*

It is worth noticing that the given algorithm is particularly effective for treating the McKean dependence of the backward component. Indeed, note that the expectation of any regular enough function of $\hat{u}$ is readily available given that the support of the approximating measure $\hat{\mu}$ coincides with the points where $\hat{u}$ is available. Of course the situation is quite different when a different approach, like a particle method, is used.

## *2.2* Main Results

In this section, we first give the rate of convergence of our algorithms 1, 2 and 3 when both the coefficients and terminal condition in (2.0.7) are smooth. This is given in Theorem 2.2.1 below. Then, we give the rate when the boundary condition is Lipschitz and when the diffusion part of (2.0.7) is uniformly non-degenerate. This does not really affect the convergence order, provided the subdivision is taken appropriately. The result is summarized in Corollary 2.2.2 below. Finally, we give the convergence of a version of our algorithm applied to equation (2.0.13) : when the dependence of the coefficients with respect to the law is general. This is given in Corollary 2.2.3.

In order to make the exposition of our results clear, let us define, for $i = 1, 2$ :

$$\mathcal{E}_u^i(k) := \max_{\pi \in \mathcal{S}_\kappa(k)} |u(T_k, \hat{X}^\pi_{T_k}) - \hat{u}^i(T_k, \hat{X}^\pi_{T_k})|; \qquad \mathcal{E}_v^i(k) := \max_{\pi \in \mathcal{S}_\kappa(k)} |v(T_k, \hat{X}^\pi_{T_k}) - \hat{v}^i(T_k, \hat{X}^\pi_{T_k})|, \tag{2.2.1}$$

with $\hat{u}^1, \hat{u}^2, \hat{v}^1$ and $\hat{v}^2$ as defined by the algorithms 2 and 3 and where $u, v$ are defined in (2.0.8), (2.0.12).

**Main result in a smooth setting.** We have that

**Theorem 2.2.1.** *Suppose that assumption* ***(SB)*** *holds. Let $m$ be a given cubature order, $q$ a given non-negative integer and $\mathcal{T}(1, q, m)$ the cubature tree defined by Algorithm 1. Then, there exists a positive constant $C$, depending only on $T$, $q$, $d$, $||\varphi_{0:d}||_{2q+m+2,\infty}$, $||\phi||_{m+2,\infty}$, such that :*

$$\max_{k \in \{0,\dots,N\}} |\langle \mu_{T_k} - \hat{\mu}_{T_k}, \phi \rangle| \le C \left(\frac{1}{N}\right)^{[(m-1)/2] \wedge q}, \tag{2.2.2}$$

*with $\hat{\mu}$ as defined in Algorithm 1.*
*Suppose in addition that $q \ge 1$ and $m \ge 3$. Then, there exists a positive constants $C_1$, depending only on $T$, $q$, $d$, $||\varphi_{0:d}||_{2q+m+2,\infty}$, $||\varphi_f||_{m+2,\infty}$,$||\phi||_{m+3,\infty}$, such that for all $k = 0, \dots, N$ :*

$$\mathcal{E}_u^1(k) + \Delta_{T_k}^{1/2} \mathcal{E}_v^1(k) \le C_1 \left(\frac{1}{N}\right), \tag{2.2.3}$$

*Moreover, suppose in addition that $q \ge 2$ and $m \ge 7$. Then, there exists a positive constant $C_2$, depending only on $T$, $q$, $d$, $||\varphi_{0:d}||_{2q+m+2,\infty}$, $||\varphi_f||_{m+2,\infty}$,$||\phi||_{m+4,\infty}$, such that for all $k = 0, \dots, N$ :*

$$\mathcal{E}_u^2(k) + \Delta_{T_k}^{1/2} \mathcal{E}_v^2(k) \le C_2 \left(\frac{1}{N}\right)^2. \tag{2.2.4}$$

**Convergence order for a Lipschitz boundary condition.**

**Corollary 2.2.2.** *Suppose that assumption* ***(LB)*** *holds. Let $m$ be a given cubature order, $q$ a given non-negative integer, $\gamma$ a non negative real and $\mathcal{T}(\gamma, q, m)$ the cubature tree defined by the algorithm 1. Then, there exists a positive constant $C$ depending only on $T$, $||\varphi||_{2q+m+2,\infty}$, $||\phi||_{1,\infty}$, such that :*

$$|\langle \mu_T - \hat{\mu}_T, \phi \rangle| \quad \le \quad C \left( \left(\frac{1}{N}\right)^{(\gamma/2) \wedge q} \vee L(\gamma, m) \right). \tag{2.2.5}$$

*where*

$$L(\gamma, m) = \begin{cases} N^{-\gamma/2} & \text{if} \quad \gamma \in (0, m-1) \\ N^{-(m-1)/2} \ln(N) & \text{if} \quad \gamma = m-1 \\ N^{-(m-1)/2} & \text{if} \quad \gamma \in (m-1, +\infty) \end{cases} \tag{2.2.6}$$

*Moreover, if $\gamma > m-1$, the results on the error control of $\hat{u}^1, \hat{v}^1; \hat{u}^2$, and $\hat{v}^2$ respectively given by* (2.2.3) *and* (2.2.4) *remain valid (with a constant $C_2'$ depending only on $T$, $q$, $d$, $||\varphi_{0:d}||_{2q+m+2,\infty}$, $||\varphi_f||_{m+2,\infty}$,$||\phi||_{1,\infty}$).*

Note that, in (2.2.5), the control holds only at time $T$ although it holds at each step in (2.2.2) : this is because the boundary condition is Lipschitz only so that we have to wait for the smoothing effect to take place.

We emphasize that the result still applies if we let the boundary condition $\phi : (y,w) \in \mathbb{R}^d \times \mathbb{R} \mapsto \phi(y,w)$ depend also on the law of the process $(X_t^x,\ 0 \le t \le T)$. For example one can consider

$$\mathbb{E}\left[\phi(X_t^x, \mathbb{E}[\varphi_\phi(X_T^x)])\right]$$

for a given $\varphi_\phi \in C_b^{m+2}$ and where $\phi$ is Lipschitz in $w$ uniformly in $y$.

The algorithm can be easily adapted to the case of the particular dependence explored in [BLP09] :

$$V_i(t,y,\mu) = \langle \mu_t, V_i(t,y,\cdot)\rangle,\ i = 0,\cdots,d,$$

and the result of Theorem 2.2.1 and Corollary 2.2.2 remain valid. Note that in that case the uniform ellipticity **(LB)** has to be understood for the matrix $[\langle \eta_t, V(t,y,\cdot)\rangle]\,[\langle \eta_t, V(t,y,\cdot)\rangle]^*$ uniformly in $y$, $t$ in $\mathbb{R}^d \times \mathbb{R}^+$ and $\eta$ family of probability measures on $\mathbb{R}^d$.

**Results under a more general law dependence.**

As mentioned in the introduction, the algorithm may be modified to solve problems in a naturally extended framework. Let us precise the framework of this extension.

Let $\mathcal{F}$ and $\mathcal{F}'$ be two classes of functions, dense in the space of continuous functions that are zero at infinity. Let $d_{\mathcal{F}}$ $d_{\mathcal{F}'}$ be two distances as defined in (2.0.14). Recall that we suppose the vector fields $V_i$ $0 \le i \le d$ that appear in (2.0.13) to be Lipschitz continuous with respect to $d_{\mathcal{F}}$ and the driver $f$ to be Lipschitz continuous with respect to $d_{\mathcal{F}'}$. Furthermore, let us suppose that there exists a unique solution $X_t^x, Y_t^x, Z_t^x$ to such a system and, as before, denote by $u$ the decoupling function defined as (2.0.8) given its existence.

Clearly, we need to modify Algorithm 1 in the natural way to be used in this framework, that is, at each discretization time, we plug directly in the coefficients the cubature based law.

In order to retrieve higher orders of convergence, we need to expand the McKean term that appears in the coefficients. In this extended case, we have to be careful when considering the forward algorithm with $q > 1$ : indeed, we must give sense to the expansion proposed at the definition of the functions $F_i$ $0 \le i \le d$ in Algorithm 1. A good notion may be the one proposed in Section 7 of [Car10]. To avoid further technicalities, we will consider here only the case with $q = 1$, i.e., when no expansion is performed.

With this definitions and observations in mind, we give the main result under a more general law dependence.

**Corollary 2.2.3.** *Let $\mu_T$ be the marginal law of the forward process in* (2.0.13) *at time $T$. Let $m \ge 3$ be a given cubature order and $\hat{\mu}_T$ be the discrete measure given by the cubature tree $\mathcal{T}(1,1,m)$ defined by the algorithm 1. Then, there exist two positive constants $C_1$ and $C_2$, depending only on $T$, $d$ such that :*

– *If **(SB)** holds and $\mathcal{F}$ (resp. $\mathcal{F}'$) is the class of functions $\varphi$ in $C_b^\infty(\mathbb{R}^d,\mathbb{R})$ (resp. $C_b^\infty(\mathbb{R}^d\times\mathbb{R},\mathbb{R})$) such that $||\varphi||_{\infty,\infty} \le 1$, then*

$$d_{\mathcal{F}}(\mu_T,\hat{\mu}_T) + \mathcal{E}_u^1(k) + \Delta_{T_k}^{1/2}\mathcal{E}_v^1(k) \le C_1 N^{-1}. \tag{2.2.7}$$

– *If **(LB)** holds and $\mathcal{F}$ (resp. $\mathcal{F}'$) is the class of functions $\varphi$ in $C_b^1(\mathbb{R}^d, \mathbb{R})$ (resp. $C_b^1(\mathbb{R}^d \times \mathbb{R}, \mathbb{R})$) such that $||\varphi||_{1,\infty} \leq 1$, then*

$$d_{\mathcal{F}}(\mu_T, \hat{\mu}_T) + \mathcal{E}_u^1(k) + \Delta_{T_k}^{1/2} \mathcal{E}_v^1(k) \leq C_2 N^{-1/2}. \tag{2.2.8}$$

We emphasize that, when $\mathcal{F} = \{\varphi \in C_b^1(\mathbb{R}^d, \mathbb{R}),\ s.t.\ ||\varphi||_{1,\infty} \leq 1\}$, thanks to the Monge-Kantorovitch duality theorem, the distance $d_{\mathcal{F}}$ is the so-called Wasserstein-1 distance.

### *2.3* Numerical examples

In this section, we illustrate the algorithm behavior by applying it to a toy model for which the exact solution is available.

Consider the $d-$dimensional MKV-FBSDE on the interval $[0, 1]$ with dynamics given by

$$\begin{aligned} dX_t &= \mathbb{E}\left[\sin(X_t)\right] dt + dB_t \\ -dY_t &= \left( \frac{\mathbf{1} \cdot \cos(X_t)}{2} + \mathbb{E}\left[ (\mathbf{1} \cdot \sin(X_t)) \exp(-Y_t^2) \right] \right) dt - Z_t \cdot dB_t, \end{aligned}$$

where $(B_t)_{0 \leq t \leq 1}$ is a $d-$dimensional Brownian motion, $\mathbf{1}$ is a $d-$dimensional vector having each entry equal to one and the sin and cos functions are applied entry-wise. Moreover, suppose that $X_0 = \mathbf{0}$. It is easily verified that a solution for the forward variable is $X = B$, and thanks to the uniqueness result this is the unique solution for the forward variable.

With respect to the backward part, take two different boundary conditions corresponding to the two considered set of assumptions **(SB)** and **(LB)**.

**(SB):** For $x \in \mathbb{R}^d$, we fix $\phi(x) = \mathbf{1} \cdot \cos(x)$. In this case, the solution to the backward part of the problem is

$$u(t, x) = \mathbf{1} \cdot \cos(x)\,; \text{ and } v(t, x) = -\sin(x),$$

which clearly implies $Y_t = \mathbf{1} \cdot \cos(X_t)$ and $Z_t = -\sin(X_t)$.

**(LB):** We fix the boundary condition to be $\phi(x) := \phi'(d^{-1/2}(\mathbf{1} \cdot x))$ where $\phi'$ is the triangular function defined for all $y \in \mathbb{R}$ as

$$\phi'(y) = \begin{cases} y + K & \text{if } y \in (-K, 0] \\ -y + K & \text{if } y \in (0, K] \\ 0 & \text{otherwise.} \end{cases}$$

In this case, the solution is given by

$$u(t, x) = \mathbb{E}\left[ \phi(X_T^x) + \int_t^T \left( \frac{\mathbf{1} \cdot \cos(X_t)}{2} + \mathbb{E}\left[ (\mathbf{1} \cdot \sin(X_t)) \exp(-Y_t^2) \right] \right) ds \right].$$

Basic properties of the Brownian motion imply that

$$u(t, x) = U(t, d^{-1/2}(\mathbf{1} \cdot x)) + (\mathbf{1} \cdot \cos(x)) \left[ \exp\left( \frac{t-1}{2} \right) - 1 \right]$$

where

$$\begin{aligned} U(t, y) =& \sqrt{\frac{1-t}{2\pi}} \left[ \exp\left( \frac{-(K+y)^2}{2(1-t)} \right) + \exp\left( \frac{-(K-y)^2}{2(1-t)} \right) - 2\exp\left( \frac{-y^2}{2(1-t)} \right) \right] \\ &+ (K+y)\left[ F\left( \frac{-y}{\sqrt{1-t}} \right) - F\left( \frac{-K-y}{\sqrt{1-t}} \right) \right] + (K-y)\left[ F\left( \frac{K-y}{\sqrt{1-t}} \right) - F\left( \frac{-y}{\sqrt{1-t}} \right) \right] \end{aligned}$$

and $F$ is the cumulative distribution function of the standard normal distribution. Evidently $v(t,x) = D_x u(t,x)$, and is defined for $t < 1$.

### *2.3.1* Tests in dimension one

Given that the law dependence already increases the dimension of the problem, we start by presenting some results when we fix $d = 1$.

***2.3.1.1 Forward component*** To implement the forward variable, we use the cubature formulae of order 3 and 5 presented in [LV04], which have paths support of size $\kappa = 2$ and $\kappa = 3$ respectively. Given the simple structure of the forward variable dynamics and the piecewise linear definition of the cubatures, we are able to solve explicitly the ODEs appearing during the tree construction. Hence there is no need to use any ODE solver.

In our first test we evaluate the weak approximation error of $X$ using the function $\phi$ as test function. Indeed we plot as error for the **(SB)** case

$$\max_{k=1,\dots,N} \left| \langle \hat{\mu}_{T_k} - \mu_{T_k}, \cos \rangle \right| ,$$

while for the **(LB)** case, we plot

$$\left| \langle \hat{\mu}_{T_k} - \mu_{T_k}, \phi \rangle \right| ,$$

where $\phi$ is the defined triangular function with $K = 0.6$. As was pointed out before, the difference between the kind of error we are observing for each case is justified as a smoothing effect is needed for the approximation to be valid under **(LB)** .

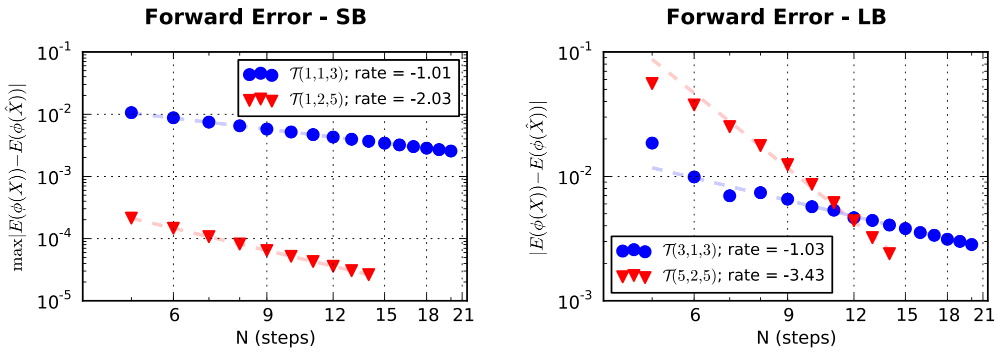


FIGURE 2. Weak approximation of the forward variable : The calculated rates are the slope of a linear regression on the last 8 points.

Figure 2 shows the obtained rate of convergence where we have used the uniform discretization in the **(SB)** case and the discretization with $\gamma = 2$ for the **(LB)** case. With the exception of the rate of convergence for the second order algorithm under **(LB)** (which is actually better than the predicted one), the expected rates of convergence are verified in both cases.

Moreover, under the smooth case, the benefit of using the higher order scheme is not only evident from a quickest convergence, but the error constant itself is smaller. This is an effect that depends on the particular example, but we remark it as it is interesting to notice that a higher order of convergence does not imply necessarily a higher initial constant.

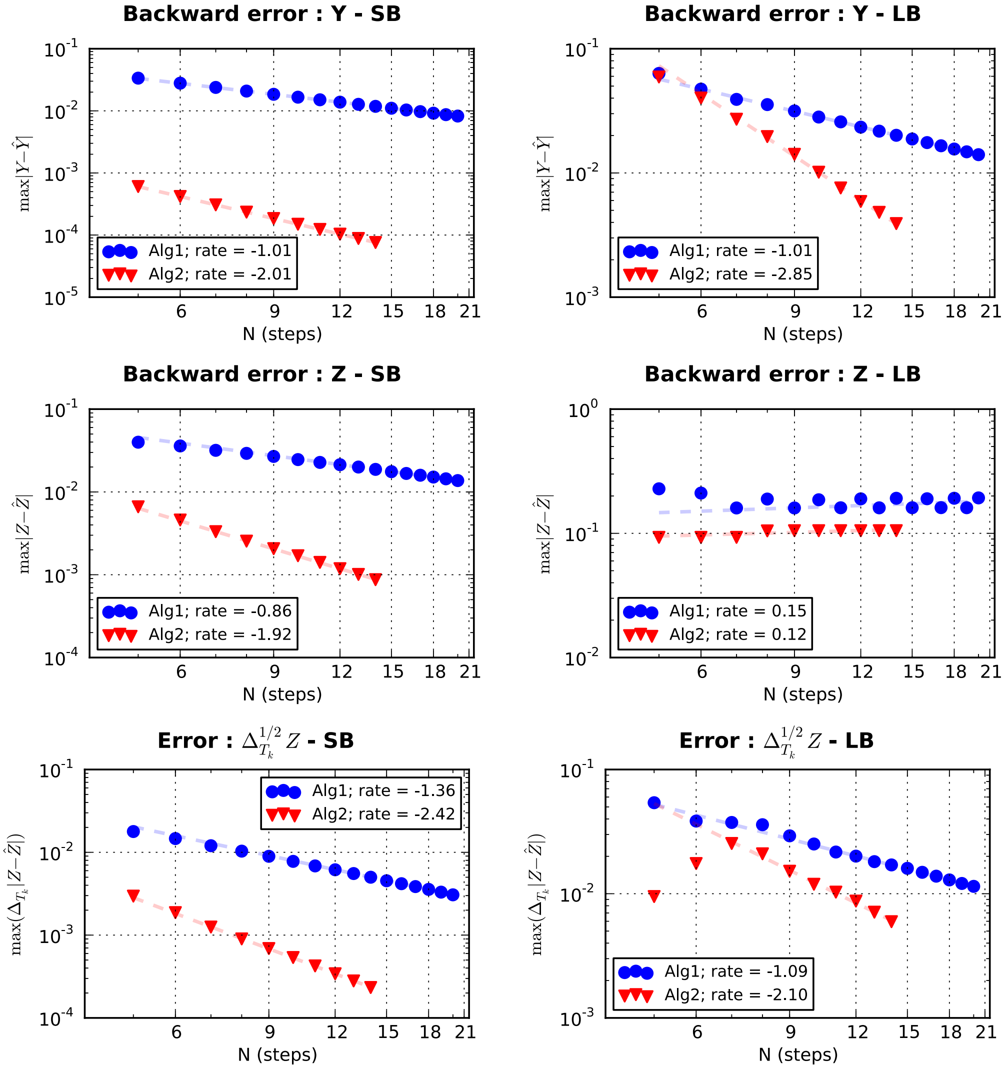


FIGURE 3. Weak approximation of the backward variable : The calculated rates are the slope of a linear regression on the last 8 points.

***2.3.1.2 Backward component*** Let us check now the approximation of the backward variable. We evaluate numerically

$$\max_{0\le k\le N-2;\ \pi\in\mathcal{S}_\kappa(k)} \left|\hat{u}^1(T_k,\hat{X}^\pi_{T_k}) - u(T_k\hat{X}^\pi_{T_k})\right| \quad\text{and}\quad \max_{0\le k\le N-2;\ \pi\in\mathcal{S}_\kappa(k)} \left|\hat{u}^2(T_k,\hat{X}^\pi_{T_k}) - u(T_k\hat{X}^\pi_{T_k})\right|;$$

for both the (SB) and (LB) cases, where we fix $K = 0.6$ for the latter.

The specific structure of our examples allows us to obtain a second order convergence scheme with a cubature of order only 5. Indeed, in such a case, the terms in front of the leading rate of convergence on the cubature error estimate (cf Claim 2.6.8) are identically 0. Given that the order 5 cubature induces a lower complexity, it is simpler to carry out simulations for a larger number of steps.

As can be appreciated from the two uppermost plots in Figure 3, the expected rates of convergence for both algorithms are verified under the smooth and Lipschitz conditions. Just as we remarked in the forward approximation, solving the backward variable in the smooth case with the higher order scheme has the double benefit of better rate of convergence and smaller constant. As one would expect, due to the use of higher order derivatives, this is no longer true for the Lipschitz case.

It is interesting to look at the behavior of the other backward variable, $Z$. We look first at an error of the type

$$\max_{0\leq k\leq N;\ \pi\in\mathcal{S}_\kappa(k)}\left|\hat{v}^1(T_k,\hat{X}^\pi_{T_k})-v(T_k\hat{X}^\pi_{T_k})\right| \quad \text{and} \quad \max_{0\leq k\leq N;\ \pi\in\mathcal{S}_\kappa(k)}\left|\hat{v}^2(T_k,\hat{X}^\pi_{T_k})-v(T_k\hat{X}^\pi_{T_k})\right|.$$

The two plots in the middle of Figure 3 are concerned with these errors. Although nice convergence is obtained in the smooth case, this is no longer true for the **(LB)** case, where the error stagnates. As will be clear from the analysis, this is a consequence of the singularity appearing at the boundary on the control of derivatives in this case. Hence, a more adequate error analysis considers errors given by

$$\max_{0\leq k\leq N;\ \pi\in\mathcal{S}_\kappa(k)}\Delta_{T_k}^{1/2}\left|\hat{v}^1(T_k,\hat{X}^\pi_{T_k})-v(T_k\hat{X}^\pi_{T_k})\right|,$$
$$\max_{0\leq k\leq N;\ \pi\in\mathcal{S}_\kappa(k)}\Delta_{T_k}^{1/2}\left|\hat{v}^2(T_k,\hat{X}^\pi_{T_k})-v(T_k\hat{X}^\pi_{T_k})\right|.$$

The expected rate of convergence of this type of error is, respectively for $\hat{v}^1$ and $\hat{v}^2$, of the same order of the order of the error of $\hat{u}^1,\hat{u}^2$ with respect to $u$. As shown in the bottommost plots in Figure 3, the numerical tests for the **(SB)** and **(LB)** cases reflect the expected rates.

### *2.3.2* Tests in higher dimensions

We evaluate as well the algorithm using our test models **(SB)** , **(LB)** with dimensions $d=2$ and $d=4$. For these tests, we evaluate only the first order schemes and use the 3-cubature formulae presented in [GL11] which have supports of size $\kappa=4$ and $\kappa=6$ respectively.

Figure 4 below shows that, just as is in the one-dimensional case, the announced rates of convergence for the forward and backward variables are verified. Note that for the particular chosen examples, the error value changes just slightly with dimension.

The case of dimensions 2 and 4 show one of the current limitations of the method : its complexity grows, in general, exponentially both in terms of the number of iterations and the dimension of the problem. Indeed, considering once again the 3-cubature formula, we have that in general the number of nodes in the tree is

$$\sharp(nodes)=\frac{(2d)^{n-1}-1}{2d-1},$$

with the obvious effects on memory management and execution time. We remark that for some particular cases, the complexity can be radically lower. For instance, under the case of constant drift and diffusion coefficients and smooth boundary conditions, using symmetric cubature formulae (as we did here) leads to a kind of "auto-pruning" of the cubature tree leading to complexity grow of the form

$$\sharp(nodes)=\sum_{i=1}^{n}i^d,$$

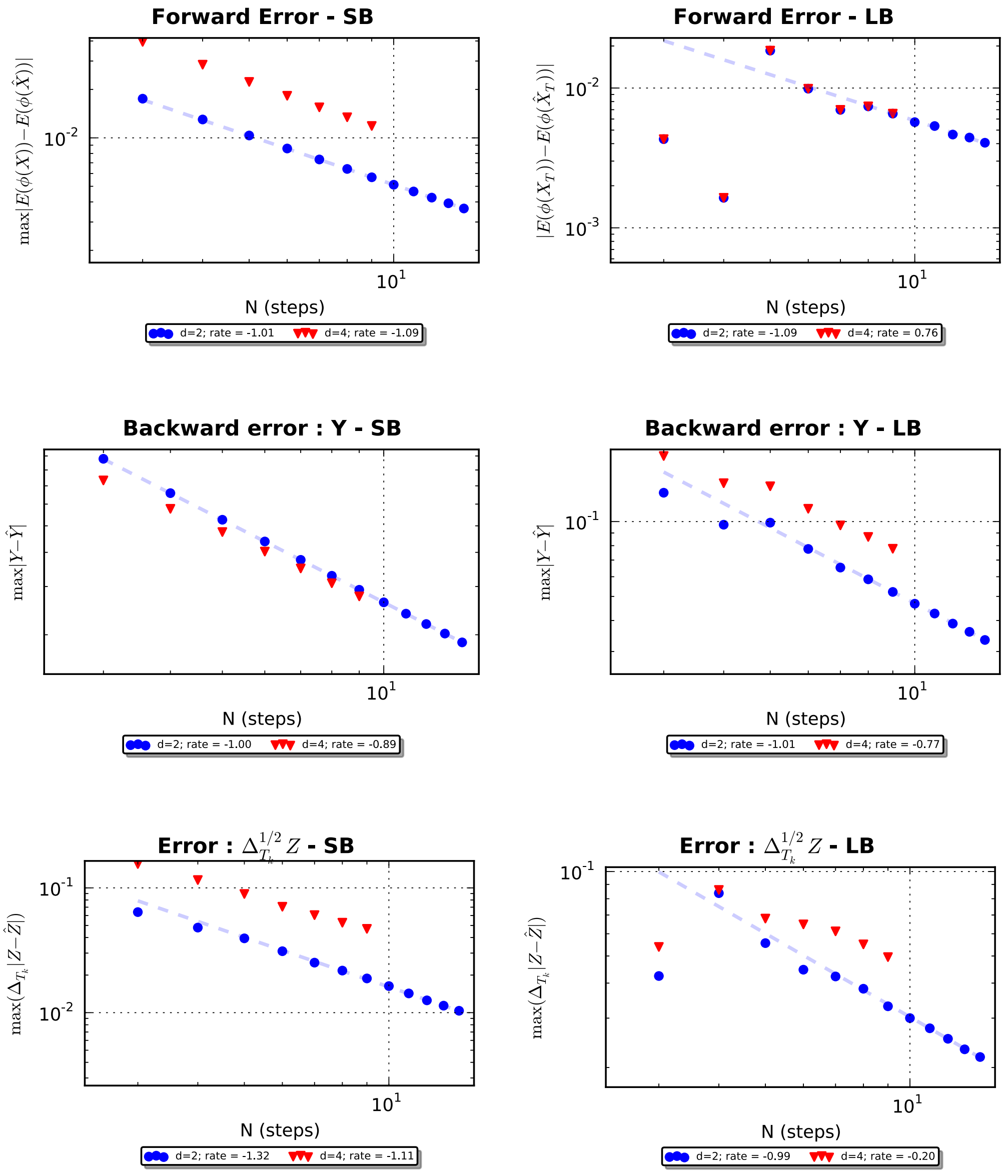


FIGURE 4. Results in dimension 2 and 4.

i.e. polynomial in $n$ with the order of the polynomial depending on the dimension $d$.

### *2.4* A class of control problems in a mean field environment

In this section, we show that equation (2.0.7) appears when solving a class of control problems inspired from the theory of mean field games but designed in such a way that the dynamics of the controlled process have no influence on the mean field environment.

For the sake of illustration, consider for instance the problem of optimization of an issuer having a large portfolio of credit assets inspired in the framework presented in [BHH+11].

One of the methods used to model credit asset dynamics is the so called structural model (see [BJPR09] for a review on credit risk models). Under this model, we assume that a credit default is triggered when the value of the corresponding credit asset is below a certain threshold. In the original Merton setup, the default may only be triggered at a certain fixed maturity time $T$. In a rather more realistic view, the default is triggered the first time the credit asset is below the threshold.

We assume that the credit assets in the basket are small and homogeneous (for example, we suppose they belong to the same economic sector) so that their value is modeled by SDEs with the same volatility and drift function terms. To simplify, we will consider the simpler Merton model. Moreover, in order to account for sector-wise contagion effects, we suppose there is a mean field type dependence in the dynamics. In addition to the credit assets, we suppose the issuer has a market portfolio used by the issuer to backup the credit risk, for example to comply with credit risk regulations, or to provide liquidity to its credit branch. Then, the value of the position of the issuer position is modeled by an SDE with coefficients depending on the contribution of all credit assets. The objective of the control problem is to maximize the value of the issuer position.

We will formalize mathematically a generalized version of the presented example. For this, we introduce a system in which a *marked* particle (the issuer in our example) with a controlled state variable $X^\alpha$ is immersed in an *environment* of $M$ interacting particles (the credit assets in our example) with state variables $X^1, \ldots, X^M$, and which dynamics are given by

$$\begin{cases} dX_t^1 = b(t, X_t^1, \frac{1}{M}\sum_{i=1}^M \delta_{X_t^i})dt + \sigma(t, X_t^1, \frac{1}{M}\sum_{i=1}^M \delta_{X_t^i})dB_t^{(1)} \\ \quad \vdots \qquad\qquad\qquad\qquad \vdots \\ dX_t^M = b(t, X_t^M, \frac{1}{M}\sum_{i=1}^M \delta_{X_t^i})dt + \sigma(t, X_t^M, \frac{1}{M}\sum_{i=1}^M \delta_{X_t^i})dB_t^{(M)} \\ dX_t^{1:M;\alpha} = b^0(t, X_t^{1:M;\alpha}, \frac{1}{M}\sum_{i=1}^M \delta_{X_t^i}, \alpha_t)dt + \sigma^0(t, X_t^{1:M;\alpha}, \frac{1}{M}\sum_{i=1}^M \delta_{X_t^i}, \alpha_t)dW_t \\ X_0^1 = \ldots X_0^M = x, \quad X_0^{1:M;\alpha} = \bar{x} \end{cases}$$

where $(\alpha_t,\ t \geq 0)$ is a progressively measurable process with image in $A \subset \mathbb{R}$, the $(B_t^{(i)})_{t\geq 0},\ i = 1, \ldots, M$ are independent $d$-dimensional Brownian motions and $(W_t,\ t \geq 0)$ is a $d$-dimensional Brownian motion. Note that, in this framework, the marked player does not influence the dynamics of the other players.

For a large number of environment players, the system is described by the McKean Vlasov system

$$\begin{cases} dX_t = b(t, X_t, \mu_t)dt + \sigma(t, X_t, \mu_t)dB_t \\ d\bar{X}_t^\alpha = b^0(t, X_t, \bar{X}_t^\alpha, \mu_t, \alpha_t)dt + \sigma^0(t, X_t, \bar{X}_t^\alpha, \mu_t)dW_t \\ X_0 = x, \quad \bar{X}_0^\alpha = \bar{x} \end{cases}$$

where $\mu_t$ is the law of $X_t$ that we will assume in the following to be fixed. Assume that the marked player is interested in minimizing the cost functional

$$J(t, x, \bar{x}, \alpha) = \mathbb{E}\left[g(X_T^{t,x}, \bar{X}_T^{\alpha;t,\bar{x}}, \mu_T) + \int_0^T f(s, X_s^{t,x}, \bar{X}_s^{\alpha;t,\bar{x}}, \mu_s)ds\right],$$

for $\alpha \in \mathcal{A}$, the set of all progressively measurable process $\alpha = (\alpha_t,\ t \geq 0)$ valued in $A$ (the maximization case is available up to a change of sign). We want to solve for the optimal value function $u(t, x, \bar{x}) = \inf\{J(t, x, \bar{x}, \alpha),\ \alpha \in \mathcal{A}\}$. Then, the associated Hamilton Jacobi

Bellman equation for $u$ reads

$$0 = D_t u(t,x,\bar{x}) + \frac{1}{2}\mathrm{Tr}(\bar{a} D^2_{x,\bar{x}} u(t,x,\bar{x})) + b(t,x,\mu_t) D_x u \\ + H(t,x,\bar{x}, D_x u, \mu_t), \tag{2.4.1}$$

with

$$\bar{a} = \begin{bmatrix} \sigma\sigma^T & \sigma\rho(\sigma^0)^T \\ \sigma^0\rho^T\sigma^T & \sigma^0(\sigma^0)^T \end{bmatrix}, \qquad \rho = [B, W]$$

and where $[\cdot,\cdot]$ stands for the quadratic variation and $H$ is the Hamiltonian

$$H(t,x,\bar{x},z,\mu_t) = \inf_{\alpha\in\mathcal{A}} \left[ b^0(t,x,\bar{x},\mu_t,\alpha) z + f(t,x,\bar{x},\mu_t) \right].$$

We will not discuss here the resolvability of the HJB equation (see e.g. [FS06] or [Pha09] for a partial revue). We can interpret (2.4.1) from a probabilistic point of view : we have that $u(t,x,\bar{x}) = Y_t^{t,x,\bar{x}}$ where $Y^{t,x,\bar{x}}$ is given by the MKV-FBSDE

$$\begin{cases} dX_s^{t,x,\bar{x}} = b(s, X_s^{t,x,\bar{x}}, \mu_s) ds + \sigma(s, X_s^{t,x,\bar{x}}, \mu_s) dB_s \\ d\bar{X}_s^{t,x,\bar{x}} = \sigma^0(s, X_s^{t,x,\bar{x}}, \bar{X}_s^{t,x,\bar{x}}, \mu_s) dW_s \\ -dY_s^{t,x,\bar{x}} = H(s, X_s^{t,x,\bar{x}}, \bar{X}_s^{t,x,\bar{x}}, \bar{Z}_s^{t,x,\bar{x}}, \mu_s) - \bar{Z}_s^{t,x,\bar{x}} dW_s + Z^{t,x,\bar{x}} dB_s \\ X_t^{t,x,\bar{x}} = x, \quad \bar{X}_t^{t,x,\bar{x}} = \bar{x}, \quad Y_0^{t,x,\bar{x}} = g(X_T^{t,x,\bar{x}}, \bar{X}_T^{t,x,\bar{x}}, \mu_T). \end{cases}$$

The reader may object that the Hamiltonian $H$ does not satisfy the boundedness condition we have assumed for the analysis of the algorithm (bounded with bounded derivatives w.r.t. the variable $z$). However, some relatively mild assumptions guarantee that the first derivative term $\bar{Z}$ will be bounded. This is almost direct when the boundary condition $g$ is bounded and smooth and proved in [CD12c] when $g$ is Lipschitz and the diffusion matrix uniformly elliptic. Hence, given an estimate on this quantity, one may introduce a modified system in which we replace in the function $(Z, \bar{Z})$ by $(\psi(Z), \bar{\psi}(\bar{Z}))$, where $\psi, \bar{\psi}$ are truncation functions used to make the value of $Z, \bar{Z}$ satisfy its known estimates, as in [Ric11] (if the estimate is not explicitly known, a sequence of functions approximating the identity may be used as in [IDR10], but some additional work would be needed to account for the truncation error). In both cases, the truncated problem will then satisfy the needed assumptions and may be solved with the presented Algorithm 1, 2.

## 2.5 Preliminaries

In the following we set a subdivision $T_0 = 0 < \cdots < T_N = T$ of $[0,T]$.

**Artificial dynamics.** We denote by $\underline{s}$ the mapping $s \mapsto \underline{s} = T_k$ if $s \in [T_k, T_{k+1})$, $k \in \{0, \cdots, N-1\}$.

For any family of probability measures $\eta^1$ and $\eta^2$, one denotes by $P^{\eta^1}$ and $\tilde{P}^{\eta^2}$ the operators such that, for all $t < s$ in $[0,T]$, for all measurable function $g$ from $\mathbb{R}^d$ to $\mathbb{R}$ and for all $y$ in $\mathbb{R}^d$ :

$$P^{\eta^1}_{t,s} g(y) = \mathbb{E}[g(X_s^{t,y,\eta^1})] \text{ and } \tilde{P}^{\eta^2}_{t,s} g(y) = \mathbb{E}[g(\tilde{X}_s^{t,y,\eta^2})]$$

and $(\mathcal{L}_s^{\eta^1})_{t\le s\le T}$ and $(\tilde{\mathcal{L}}_s^{\eta^2})_{t\le s\le T}$ their infinitesimal generator, where for all $g$ in $C^2(\mathbb{R}^d, \mathbb{R})$

$$\mathcal{L}_s^{\eta^1} g(y) := V_0(s, y, \langle \eta^1_s, \varphi_0\rangle) \cdot D_y g(y) + \frac{1}{2}\mathrm{Tr}[VV^T(s,y,\langle\eta^1_s,\varphi\rangle) D_y^2 g(y)] \tag{2.5.1}$$

and by definition $\tilde{\mathcal{L}}_s^{\eta^2} = \mathcal{L}_{\underline{s}}^{\eta^2}$. Here $X^{t,y,\eta^1}$ and $\tilde{X}^{t,y,\eta^2}$ are the respective solutions of

$$dX_s^{t,y,\eta^1} = \sum_{i=0}^{d} V_i\left(s, X_s^{t,y,\eta^1}, \langle \eta_s^1, \varphi_i\rangle\right) dB_s^i, \quad X_t^{t,y,\eta^1} = y, \tag{2.5.2}$$

$$d\tilde{X}_s^{t,y,\eta^2} = \sum_{i=0}^{d} V_i\left(s, \tilde{X}_s^{t,y,\eta^2}, \sum_{p=0}^{q-1}[(t-\underline{t})^p/p!]\langle \eta_{\underline{s}}^2, (\tilde{\mathcal{L}}^{\eta^2})^p\varphi_i\rangle\right) dB_t^i, \quad \tilde{X}_t^{t,y,\eta^2} = y. \tag{2.5.3}$$

Finally, let us define the operator associated to the cubature measure, $Q^{\hat{\mu}}$, as

$$Q_{t,s}^{\hat{\mu}} g(y) = \mathbb{E}_{\mathbb{Q}_{t,s}}[g(\tilde{X}_s^{t,y,\hat{\mu}})] \tag{2.5.4}$$

for all $t < s$ in $[0,T]$ for all $y$ in $\mathbb{R}^d$ and for all measurable function $g$ from $\mathbb{R}^d$ to $\mathbb{R}$. Note that for all $k$ in $\{1,\dots,N\}$ :

$$Q_{0,T_k}^{\hat{\mu}} g(x) = \langle \hat{\mu}_{T_k}, g\rangle.$$

**Multi-index (2).** Let $\mathcal{M}$ be defined by (2.1.1). Let $\beta \in \mathcal{M}$. We define $|\beta| = l$ if $\beta = (\beta_1,\dots,\beta_l)$, $|\beta|_0 := \mathrm{card}\{i : \beta_i = 0\}$ and $\|\beta\| := |\beta| + |\beta|_0$. Naturally $|\emptyset| = |\emptyset|_0 = \|\emptyset\| = 0$. For every $\beta \neq \emptyset$, we set $-\beta := (\beta_2,\dots,\beta_l)$ and $\beta- := (\beta_1,\dots,\beta_{l-1})$. We set $\beta^+$ the multi-index obtained by deleting the zero components of $\beta$.

We will frequently refer to the set of multi-indices of degree at most $l$ denoted by $\mathcal{A}_l := \{\beta \in \mathcal{M} : \|\beta\| \leq l\}$. We define as well its *frontier set* $\partial\mathcal{A} := \{\beta \in \mathcal{M} \setminus \mathcal{A} : -\beta \in \mathcal{A}\}$. We can easily check that $\partial\mathcal{A}_l \subset \mathcal{A}_{l+2} \setminus \mathcal{A}_l$.

**Directional derivatives.** For notational convenience, let us define the second order operator

$$\mathcal{V}_{(0)} := \partial_t + \mathcal{L},$$

and for $j = \{1,\dots,d\}$ the operator

$$\mathcal{V}_{(j)} := V_j \frac{\partial}{\partial x_j}.$$

where, as announced in the notation section, we do not mark explicitly the time, space and law dependence. For every $\|\beta\| \leq l$ let us define recursively

$$\mathcal{V}_\beta g := \begin{cases} g & \text{if } |\beta| = 0 \\ V_{\beta_1}\mathcal{V}_{-\beta} g & \text{if } |\beta| > 0, \end{cases} \tag{2.5.5}$$

provided that $g : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}$ is smooth enough. Hence, for $n \in \mathbb{N}$ we denote by $\mathcal{D}_b^n$, the space of such functions $g$ for which $\mathcal{V}_\beta g$ exists and is bounded for every $\beta \in \mathcal{A}_n$. For any function $g$ in $\mathcal{D}_b^n$, we set for all $\beta \in \mathcal{A}_n$,

$$D_\beta g := \frac{\partial}{\partial y_{\beta_1}} \cdots \frac{\partial}{\partial y_{\beta_{|\beta|}}} g,$$

where $[\partial/\partial y_0]$ must be understood as $[\partial/\partial t]$.

**Iterated integrals.** For any multi-index $\beta$ and adapted process $g$ we define for all $t < s \in [0,T]$ the *multiple Itô integral* $I_\beta^{t,s}[g]$ recursively by

$$I_\beta^{t,s}(g) = \begin{cases} g(\tau) & \text{if } |\beta| = 0 \\ \int_t^s I_{\beta-}^{\rho,r}(g)dr & \text{if } |\beta| > 0 \text{ and } \beta_l = 0 \\ \int_t^s I_{\beta-}^{t,r}(g)dB_r^{\beta_l} & \text{if } |\beta| > 0 \text{ and } \beta_l > 0 \end{cases}$$

We will write $I_\beta^{t,s} := I_\beta^{t,s}(1)$.

The previous notation is very convenient to introduce an Itô-Taylor expansion, that is an analogue of Taylor formula when dealing with Itô processes. The proof follows simply by repeated iteration of Itô's lemma, and may be found (without the law dependence) in [KP92].

**Lemma 2.5.1.** *Let $t < s \in [0,T]$ and $y \in \mathbb{R}^d$. Let $n \in \mathbb{N}^*$ and let $g$ in $D_b^n$ Then, for each family of probability measures $\eta$ on $\mathbb{R}^d$, we have have an* Itô-Taylor expansion of order $n$, *that is*

$$g(s, X_s^{t,y,\eta}) = g(t,y) + \sum_{\beta \in \mathcal{A}_n} \mathcal{V}_\beta g(t,y) I_\beta^{t,s} + \sum_{\beta \in \partial\mathcal{A}_n} I_\beta^{t,s}[\mathcal{V}_\beta g(., X_.^{t,y,\eta})]$$

*where $(X_s^{t,y,\eta},\ t \leq s \leq T)$ is the solution of* (2.5.3).

The following lemma is a particular case of a result in [KP92]. It follows from integration by parts formula and expectation properties.

**Lemma 2.5.2.** *Let $\beta \in \mathcal{M}$, and let $t_1 < t_2 \in [0,T]$. Then for any bounded and measurable functions $g_1$ and $g_2$ in $[t_1,t_2]$ there exists a constant depending only on $\beta$ and $i$ such that*

$$\mathbb{E}\left[I_\beta^{t_1,t_2}(g_1) I_{(i)}^{t_1,t_2}(g_2) \middle| \mathcal{F}_{t_1}\right] \leq \mathbf{1}_{\{(\beta)^+ = i\}} C(\beta,i)(t_2-t_1)^{(||\beta||+1)/2} \sup_{t_1 \leq s \leq t_2} |g_1(s)| \sup_{t_1 \leq s \leq t_2} |g_2(s)|,$$

$$\mathbb{E}\left[I_\beta^{t_1,t_2}(g_1) I_{(0,i)}^{t_1,t_2}(g_2) \middle| \mathcal{F}_{t_1}\right] \leq \mathbf{1}_{\{(\beta)^+ = i\}} C'(\beta,i)(t_2-t_1)^{(||\beta||+3)/2} \sup_{t_1 \leq s \leq t_2} |g_1(s)| \sup_{t_1 \leq s \leq t_2} |g_2(s)|.$$

## *2.6* Proof of Theorem 2.2.1 under (SB)

### *2.6.1* Rate of convergence of the forward approximation : proof of (2.2.2)

Here we prove the approximation order of the forward component. Let $\eta$ be a given family of probability measures on $\mathbb{R}^d$. We have the following decomposition of the error :

$$\begin{aligned}
(P_{T_0,T_N} - Q_{T_0,T_N}^{\hat\mu})\phi(x) &= (P_{T_0,T_N} - P_{T_0,T_N}^{\eta})\phi(x) + Q_{T_0,T_{N-1}}^{\hat\mu}(P_{T_{N-1},T_N}^{\eta} - Q_{T_{N-1},T_N}^{\hat\mu})\phi(x) \\
&\quad + Q_{T_0,T_{N-2}}^{\hat\mu} P_{T_{N-2},T_N}^{\eta}\phi(x) - Q_{T_0,T_{N-1}}^{\hat\mu} P_{T_{N-1},T_N}^{\eta}\phi(x) \\
&\quad + Q_{T_0,T_{N-3}}^{\hat\mu} P_{T_{N-3},T_N}^{\eta}\phi(x) - Q_{T_0,T_{N-2}}^{\hat\mu} P_{T_{N-2},T_N}^{\eta}\phi(x) \\
&\quad \vdots \\
&\quad + P_{T_0,T_N}^{\eta}\phi(x) - Q_{T_0,T_1}^{\hat\mu} P_{T_1,T_N}^{\eta}\phi(x) \\
&= (P_{T_0,T_N} - P_{T_0,T_N}^{\eta})\phi(x) + (P_{T_0,T_1}^{\eta} - Q_{T_0,T_1}^{\hat\mu}) P_{T_1,T_N}^{\eta}\phi(x) \\
&\quad + \sum_{j=1}^{N-1} Q_{T_0,T_j}^{\hat\mu}\left[(P_{T_j,T_{j+1}}^{\eta} - Q_{T_j,T_{j+1}}^{\hat\mu}) P_{T_{j+1},T_N}^{\eta}\phi(x)\right],
\end{aligned}$$

so that :

$$\begin{aligned}(P_{T_0,T_N}-Q^{\hat{\mu}}_{T_0,T_N})\phi(x) &= (P_{T_0,T_N}-P^{\eta}_{T_0,T_N})\phi(x)+(P^{\eta}_{T_0,T_1}-Q^{\hat{\mu}}_{T_0,T_1})P^{\eta}_{T_1,T_N}\phi(x)\\ &+\sum_{j=1}^{N-2}Q^{\hat{\mu}}_{T_0,T_j}\left[(P^{\eta}_{T_j,T_{j+1}}-Q^{\hat{\mu}}_{T_j,T_{j+1}})P^{\eta}_{T_{j+1},T_N}\phi(x)\right]\\ &+Q^{\hat{\mu}}_{T_0,T_{N-1}}\left[(P^{\eta}_{T_{N-1},T_N}-Q^{\hat{\mu}}_{T_{N-1},T_N})\phi(x)\right]\end{aligned}$$

That is, the global error is decomposed as a sum of local errors (i.e., as the sum of errors on each interval). These local errors can be also split. Let us define the function

$$\psi(T_k,x) := P^{\eta}_{T_k,T_N}\phi(x). \tag{2.6.1}$$

We emphasize that for all $t$ in $[0,T)$, $y\mapsto\psi(t,y)$ is $C_b^\infty$. Indeed, this function can be seen as the solution of the PDE

$$\begin{cases}\partial_t\psi(t,y)+\mathcal{L}^{\eta}\psi(t,y)=0, \text{ on } [0,T]\times\mathbb{R}^d\\ \psi(T,y)=\phi(y)\end{cases} \tag{2.6.2}$$

taken at time $T_k$, where $\mathcal{L}^\eta$ is defined by (2.5.1). The claim follows from Lemma 2.7.1. On each interval $\Delta_{T_k}$, $k=1,\cdots,N-1$, the local error $(P^{\eta}_{T_k,T_{k+1}}-Q^{\hat{\mu}}_{T_k,T_{k+1}})\psi(T_{k+1},x)$ is expressed as

$$\begin{aligned}(P^{\eta}_{T_k,T_{k+1}}-Q^{\hat{\mu}}_{T_k,T_{k+1}})\psi(T_{k+1},x) &= (P^{\eta}_{T_k,T_{k+1}}-\tilde{P}^{\hat{\mu}}_{T_k,T_{k+1}})\psi(T_{k+1},x) && (2.6.3)\\ &+(\tilde{P}^{\hat{\mu}}_{T_k,T_{k+1}}-Q^{\hat{\mu}}_{T_k,T_{k+1}})\psi(T_{k+1},x). && (2.6.4)\end{aligned}$$

Error (2.6.3) can be identified as a frozen (in time) error (and so, a sort of weak Euler error) plus an approximation error, in the sense that in step $k$, the measure $\mu_{T_k}$ is approximated by the discrete law $\hat{\mu}_{T_k}$. Then, (2.6.4) is a (purely) cubature error on one step, and we have :

$$\begin{aligned}&(P_{T_0,T_N}-Q^{\hat{\mu}}_{T_0,T_N})\phi(x)\\ &=(P_{T_0,T_N}-P^{\eta}_{T_0,T_N})\phi(x)\\ &\quad+(P^{\eta}_{T_0,T_1}-\tilde{P}^{\hat{\mu}}_{T_0,T_1})\psi(T_1,x)+\sum_{j=1}^{N-2}Q^{\hat{\mu}}_{T_0,T_j}\left[(P^{\eta}_{T_j,T_{j+1}}-\tilde{P}^{\hat{\mu}}_{T_j,T_{j+1}})\psi(T_{j+1},x)\right]\\ &\quad+(\tilde{P}^{\hat{\mu}}_{T_0,T_1}-Q^{\hat{\mu}}_{T_0,T_1})\psi(T_1,x)+\sum_{j=1}^{N-2}Q^{\hat{\mu}}_{T_0,T_j}\left[(\tilde{P}^{\hat{\mu}}_{T_j,T_{j+1}}-Q^{\hat{\mu}}_{T_j,T_{j+1}})\psi(T_{j+1},x)\right]\\ &\quad+Q^{\hat{\mu}}_{T_0,T_{N-1}}\left[(P^{\eta}_{T_{N-1},T_N}-Q^{\hat{\mu}}_{T_{N-1},T_N})\phi(x)\right]\end{aligned}$$

We have the two following Claims :

**Claim 2.6.1.** *There exists a positive constant $C(T,V_{0:d})$ depending on the regularity of the $V_{0:d}$ and on $T$ such that, for all $y$ in $\mathbb{R}^d$, for all family of probability measures $\eta$, for all*

$k \in \{1, \cdots, N-1\}$, *one has :*

$$
\begin{aligned}
&\left|(P^{\eta}_{T_k,T_{k+1}} - \tilde{P}^{\hat{\mu}}_{T_k,T_{k+1}})\psi(T_{k+1},y)\right| \\
&\leq C(T,V_{0:d})||\psi(T_{k+1},.)||_{2,\infty} \sum_{i=0}^{d} \int_{T_k}^{T_{k+1}} \left|\langle \eta_t, \varphi_i\rangle - \sum_{p=0}^{q-1}[(t-T_k)^p/p!]\langle \eta_{T_k}, (\mathcal{L}^{\eta})^p\varphi_i\rangle\right| dt \\
&\quad + C(T,V_{0:d})||\psi(T_{k+1},.)||_{2,\infty} \sum_{i=0}^{d}\sum_{p=0}^{q-1}[\Delta^p_{T_{k+1}}/p!]\left|\langle \eta_{T_k}, (\mathcal{L}^{\eta})^p\varphi_i\rangle - \langle \hat{\mu}_{T_k}, (\mathcal{L}^{\hat{\mu}})^p\varphi_i\rangle\right|.
\end{aligned}
$$

*Proof.* We deduce the claim by applying Lemma 2.7.3 to $y \mapsto \psi(T_k, y)$ for each $k \in \{0, \cdots, N-2\}$. $\square$

**Claim 2.6.2.** *There exists a positive constant $C(T, V_{0:d}, d, m)$ depending on the regularity of the $V_{0:d}$ on the dimension d and the cubature order m such that : for all y in $\mathbb{R}^d$, for all family of probability measure $\eta$, for all $k \in \{0, \cdots, N-1\}$, one has :*

$$
\left|(\tilde{P}^{\hat{\mu}}_{T_k,T_{k+1}} - Q^{\hat{\mu}}_{T_k,T_{k+1}})\psi(T_{k+1},y)\right| \leq C(T,V_{0:d},d,m)\sum_{l=m+1}^{m+2}||\psi(T_{k+1},\cdot)||_{l,\infty}\Delta^{\frac{l}{2}}_{T_{k+1}}.
$$

*Proof.* The claim follows, by applying Lemma 2.7.4 with $\hat{\mu}$ to the function $y \in \mathbb{R}^d \mapsto \psi(T_k, y)$ for each $k \in \{1, \cdots, N-1\}$. $\square$

Then, by plugging estimates of Claims 2.6.1 and 2.6.2 in the error expansion we deduce that :

$$
\begin{aligned}
&\left|(P_{T_0,T_N} - Q^{\hat{\mu}}_{T_0,T_N})\phi(x)\right| \qquad (2.6.5)\\
&\leq C(T,V_{0:d})\sum_{j=0}^{N-1}||\psi(T_{j+1},\cdot)||_{2,\infty}\int_{T_j}^{T_{j+1}}\sum_{i=0}^{d}\left|\langle \eta_t,\varphi_i\rangle - \sum_{p=0}^{q-1}[(t-T_k)^p/p!]\langle \eta_{T_j},(\mathcal{L}^{\eta})^p\varphi_i\rangle\right| dt \\
&\quad + C(T,V_{0:d})\sum_{j=0}^{N-1}||\psi(T_{j+1},\cdot)||_{2,\infty}\sum_{p=0}^{q-1}\Delta^{p+1}_{T_{j+1}}\sum_{i=0}^{d}\left|\langle \eta_{T_j},(\mathcal{L}^{\eta})^p\varphi_i\rangle - \langle \hat{\mu}_{T_j},(\mathcal{L}^{\hat{\mu}})^p\varphi_i\rangle\right| \\
&\quad + C(T,V_{0:d},d,m)\sum_{j=0}^{N-1}\sum_{l=m+1}^{m+2}||\psi(T_{j+1},\cdot)||_{l,\infty}\Delta^{\frac{l}{2}}_{T_{j+1}} + \left|(P_{T_0,T_N} - P^{\eta}_{T_0,T_N})\phi(x)\right|
\end{aligned}
$$

Up to now, the analysis holds for any family of probability measures $\eta$. *The key point in the proof is to note that we can actually choose $\eta = \mu$*, that is, the law of the solution of the forward component in (2.0.7). In that case for all measurable function $g$ :

$$
\langle \eta_., g\rangle = \langle \mu_., g\rangle = \mathbb{E}[g(X^x_.)].
$$

Then, (2.6.5) becomes :

$$
\begin{aligned}
&\left|(P_{T_0,T_N} - Q^{\hat{\mu}}_{T_0,T_N})\phi(x)\right| \\
&\leq C(T,V_{0:d}) \sum_{j=0}^{N-1} ||\psi(T_{j+1},\cdot)||_{2,\infty} \int_{T_j}^{T_{j+1}} \sum_{i=0}^{d} \Big|\mathbb{E}\left[\varphi_i(X_t^{x,\mu})\right] \\
&\quad - \sum_{p=0}^{q-1}[(t-T_k)^p/p!]\mathbb{E}\left[(\mathcal{L}^{\mu})^p\varphi_i(X_{T_j}^{x,\mu})\right]\Big| \, dt \\
&\quad + C(T,V_{0:d}) \sum_{j=0}^{N-1} ||\psi(T_{j+1},\cdot)||_{2,\infty} \sum_{p=0}^{q-1} \Delta_{T_{j+1}}^{p+1} \sum_{i=0}^{d} \left|\mathbb{E}\left[(\mathcal{L}^{\mu})^p\varphi_i(X_{T_j}^{x,\mu})\right] - \mathbb{E}\left[(\mathcal{L}^{\hat{\mu}})^p\varphi_i(\hat{X}_{T_j}^{x,\hat{\mu}})\right]\right| \\
&\quad + C(T,V_{0:d},d,m) \sum_{j=0}^{N-1} \sum_{l=m+1}^{m+2} ||\psi(T_{j+1},\cdot)||_{l,\infty} \Delta_{T_{j+1}}^{\frac{l}{2}}
\end{aligned}
\tag{2.6.6}
$$

since $P_{s,t} = P^{\mu}_{s,t}$ for all $s < t \in [0,T]$, by definition. Now we have :

**Claim 2.6.3.** *For any $k \in \{0,\cdots,N-1\}$, and for all $t$ in $[T_k;T_{k+1})$ there exists a positive constant $C(d,V_{0:d})$ such that :*

$$\int_{T_k}^{T_{k+1}} \sum_{i=0}^{d} \left|\mathbb{E}\left[\varphi_i(X_t^{x,\mu})\right] - \sum_{p=0}^{q-1}[(t-T_k)^p/p!]\mathbb{E}\left[(\mathcal{L}^{\mu})^p\varphi_i(X_{T_k}^{x,\mu})\right]\right| dt \leq C(d,V_{0:d})\|\varphi\|_{2q,\infty}\Delta_{T_{k+1}}^{q+1}$$

*Proof.* This follows by Itô-Taylor expansion of order $q$ of $\varphi_i(X^{X_{T_k},\mu})$ for each $i = 0,\cdots,d$ and for any $k$ in $\{0,\cdots,N-1\}$. ◻

Therefore,

$$
\begin{aligned}
&\left|(P_{T_0,T_N} - Q^{\hat{\mu}}_{T_0,T_N})\phi(x)\right| \\
&\leq C(T,V_{0:d}) \sum_{j=0}^{N-1} ||\psi(T_{j+1},\cdot)||_{2,\infty} \sum_{p=0}^{q-1} \Delta_{T_{j+1}}^{p+1} \sum_{i=0}^{d} \left|(P_{T_0,T_N} - Q^{\hat{\mu}}_{T_0,T_N})\mathcal{L}^p\varphi_i(x)\right| \\
&\quad + C(T,V_{0:d},d)\|\varphi\|_{2q,\infty} \sum_{j=0}^{N-1} ||\psi(T_{j+1},\cdot)||_{2,\infty}\Delta_{T_{j+1}}^{q+1} \\
&\quad + C(T,V_{0:d},d,m) \sum_{j=0}^{N-1} \sum_{l=m+1}^{m+2} ||\psi(T_{j+1},\cdot)||_{l,\infty} \Delta_{T_{j+1}}^{\frac{l}{2}}.
\end{aligned}
\tag{2.6.7}
$$

Thanks to estimate (2.7.2) in Lemma 2.7.1, for all $n$ in $\mathbb{N}$, we have the following bound on the supremum norm of the derivatives of $\psi$ up to order $n$ :

$$||\nabla_y^n\psi(t,\cdot)||_{\infty} \leq C(T,V_{0:d})||\phi||_{n,\infty}. \tag{2.6.8}$$

By plugging this bound in (2.6.7) we get

$$\begin{aligned}
&\left|(P_{T_0,T_N}-Q^{\hat{\mu}}_{T_0,T_N})\phi(x)\right|\\
&\leq C(T,V_{0:d})||\phi||_{2,\infty}\sum_{j=0}^{N-1}\sum_{p=0}^{q-1}\Delta_{T_{j+1}}^{p+1}\sum_{i=0}^{d}\left|(P_{T_0,T_j}-Q^{\hat{\mu}}_{T_0,T_j})\mathcal{L}^p\varphi_i(x)\right|\\
&\quad +C(T,V_{0:d},d,m)(||\phi||_{2m+2,\infty}+||\varphi||_{2q,\infty})\left(\frac{1}{N}\right)^{q\wedge(m-1)/2}.
\end{aligned}$$

It should be remarked that the term on the right hand side is controlled in terms of the approximation error itself, acting on the functions $\mathcal{L}^p\varphi_i,\ i=1,\cdots,d,\ p=0,\cdots,q-1$, from step 0 to $j$ for any $j$ in $\{1,\cdots,N-1\}$. To proceed, the argument is the following one : since these bounds hold for all (at least) $\phi$ smooth enough, one can let $\phi=\varphi_0$, and use the discrete Gronwall Lemma to get the following bound on $\varphi_0$ :

$$\begin{aligned}
&\left|(P_{T_0,T_N}-Q^{\hat{\mu}}_{T_0,T_N})\varphi_0(x)\right|\\
&\leq C(T,V_{0:d},m,||\phi||_{m+2,\infty},||\varphi||_{2q,\infty},||\varphi_0||_{2q+m,\infty})\left(\frac{1}{N}\right)^{q\wedge[(m-1)/2]}\\
&\quad\times\Bigg\{\sum_{j=0}^{N-1}\sum_{p=0}^{q-1}\Delta_{T_{j+1}}^{p+1}\sum_{i=1}^{d}\left|(P_{T_0,T_j}-Q^{\hat{\mu}}_{T_0,T_j})\mathcal{L}^p\varphi_i(x)\right|\\
&\quad\quad+\sum_{j=0}^{N-1}\sum_{p=1}^{q-1}\Delta_{T_{j+1}}^{p+1}\left|(P_{T_0,T_j}-Q^{\hat{\mu}}_{T_0,T_j})\mathcal{L}^p\varphi_0(x)\right|\Bigg\}
\end{aligned}$$

It is clear that, by iterating this argument (i.e., by letting $\phi=\varphi_1$ and then $\phi=\varphi_2$,..., $\phi=\mathcal{L}\varphi_0$, etc...) we obtain :

$$\left|(P_{T_0,T_N}-Q^{\hat{\mu}}_{T_0,T_N})\phi(x)\right|\quad\leq\quad C(T,V,d,q,m,||\phi||_{m+2,\infty},||\varphi||_{2q+m,\infty})\left(\frac{1}{N}\right)^{q\wedge[(m-1)/2]}.$$

This concludes the proof (2.2.2) at time $T$. From these arguments, we easily deduce that the estimate holds for any $T_k$, $k=1,\ldots,N$. □

### *2.6.2* Rate of convergence for the backward approximation : proof of (2.2.3) and (2.2.4)

Here we prove the approximation order of the backward component. Before presenting the proof, we introduce some notations. Let us define the Brownian counterparts of $\hat{\Theta}_{k+1,k},\hat{\Theta}_k$ and $\hat{\zeta}_k$ given in step 9 in Algorithm 2 and steps 12, 18 and 13 in Algorithm 3. For all family of probability measures $\eta$ we set

$$\begin{aligned}
\Theta_k^{\eta}(y) &:= (T_l,y,u(T_k,y),v(T_k,y),\langle\eta_{T_k},\varphi_f[\cdot,u(T_k,\cdot)]\rangle)\,, &(2.6.9)\\
\bar{\Theta}_{k+1,k}^{\eta^1,\eta^2}(y) &:= \left(T_{k+1},X_{T_{k+1}}^{T_k,y,\eta^1},u(T_{k+1},X_{T_{k+1}}^{T_k,y,\eta^1}),v(T_k,y),\langle\eta^2_{T_k},\varphi_f[\cdot,u(T_k,\cdot)]\rangle\right),
\end{aligned}$$

and

$$\zeta_k=4\frac{B_{T_{k+1}}-B_{T_k}}{\Delta_{T_{k+1}}}-6\frac{\int_{T_k}^{T_{k+1}}(s-T_k)dB_s}{\Delta^2_{T_{k+1}}}.$$

The proof uses extensively the regularity of the function $u$. From Lemma 2.7.2, for all $t\in[0,T)$, the function $y\in\mathbb{R}^d\mapsto u(t,y)$ is $C_b^\infty$ with uniform bounds in time. In the elliptic case the same situation holds, although the bounds depend on time and blow up in the boundary. Hence, we keep track of the explicit dependence of each error term on $u$ and its derivatives in such a way that the proof is simplified for the elliptic case.

Moreover, we will expand and bound terms of the form $y \mapsto f(\cdot, y, u(\cdot, y), \mathcal{V}u(\cdot, y), \cdot)$. When differentiating such a term, the bounds involve the product of the derivatives of $u$ with respect to the space variable. Namely, the $r^{\text{th}}$ differentiation of $f$ involves a product of at most $r+1$ derivatives of $u$. To keep track of the order of the derivatives that appear in the bound, we introduce the set of positive integers for which their sum is less than or equal to $r$ : $\mathcal{I}(l,r) = \{I = (I_1, \ldots, I_l) \in \{1, \ldots, r\}^l : \sum_j I_j \le r\}$ and the following quantity :

$$M_u(r,s) = \sum_{l=1}^{r} \sum_{I \in \mathcal{I}(l,r)} \prod_{j=1}^{l} ||u(s,.)||_{I_j,\infty}. \tag{2.6.10}$$

*(1) Proof of the order of convergence for the first order algorithm (Algorithm 2).*
Let $k \in \{1, \ldots, N-1\}$. We first break the error between $u$ and $\hat{u}^1$ as follows :

$$\begin{aligned}
&u(T_k, \hat{X}^\pi_{T_k}) - \hat{u}^1(T_k, \hat{X}^\pi_{T_k}) \\
&\quad = u(T_k, \hat{X}^\pi_{T_k}) - \mathbb{E}\left[u(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \mu}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\bar{\Theta}^{\mu,\mu}_{k+1,k}(\hat{X}^\pi_{T_k}))\right] && (2.6.11)\\
&\quad +\mathbb{E}\left[u(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \mu}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\bar{\Theta}^{\mu,\mu}_{k+1,k}(\hat{X}^\pi_{T_k}))\right] && (2.6.12)\\
&\qquad - \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\left[u(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \hat{\mu}}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\bar{\Theta}^{\hat{\mu},\mu}_{k+1,k}(\hat{X}^\pi_{T_k}))\right] \\
&\quad +\mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\Bigg[u(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \hat{\mu}}_{T_{k+1}}) - \hat{u}^1(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \hat{\mu}}_{T_{k+1}}) && (2.6.13)\\
&\qquad + \Delta_{T_{k+1}}\left(f(\bar{\Theta}^{\hat{\mu},\mu}_{k+1,k}(\hat{X}^\pi_{T_k})) - f(\hat{\Theta}^{\pi,1}_{k+1,k})\right)\Bigg].
\end{aligned}$$

Similarly, we can expand the error between $v$ and $\hat{v}^1$ as :

$$\begin{aligned}
&\Delta_{T_{k+1}}\left[v(T_k, \hat{X}^\pi_{T_k}) - \hat{v}^1(T_k, \hat{X}^\pi_{T_k})\right] \\
&= \Delta_{T_{k+1}} v(T_k, \hat{X}^\pi_{T_k}) - \mathbb{E}\left[u\left(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \mu}_{T_{k+1}}\right)\Delta B_{T_{k+1}}\right] && (2.6.14)\\
&\quad +\mathbb{E}\left[u\left(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \mu}_{T_{k+1}}\right)\Delta B_{T_{k+1}}\right] \\
&\qquad -\mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\left[u\left(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \hat{\mu}}_{T_{k+1}}\right)\Delta B_{T_{k+1}}\right] && (2.6.15)\\
&\quad +\mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\left(\left[u\left(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \hat{\mu}}_{T_{k+1}}\right) - \hat{u}^1\left(T_{k+1}, X^{T_k, \hat{X}^\pi_{T_k}, \hat{\mu}}_{T_{k+1}}\right)\right]\Delta B_{T_{k+1}}\right) && (2.6.16)
\end{aligned}$$

Then, at each step, the approximation error on the backward variables can be expanded as : a first term (2.6.11) and (2.6.14), corresponding to scheme errors ; a second term, (2.6.12) and (2.6.15), corresponding to generalized cubature errors and can be viewed as one step versions of the forward error (2.2.2) in Theorem 2.2.1 ; and a third term, (2.6.13) and (2.6.16), which are propagation errors.

Let us explain how the proof works. We will bound separately each error : the scheme, cubature and propagation errors. Each bound is summarized in a Claim (respectively Claims

2.6.4, 2.6.5 and 2.6.5 below). Then, we will deduce the dynamics of the error at step $k$, $\mathcal{E}_u^1(k)$ defined as (2.2.1) and conclude with a Gronwall argument.

The first claim below gives the bounds on the scheme errors.

**Claim 2.6.4.** *There exists a constant $C$ depending on the regularity of $V_{0:d}$ and $f$ (and not on $k$) such that the scheme errors (2.6.11) and (2.6.14) are bounded by :*

$$\begin{aligned}
&\left|u(T_k,\hat{X}^{\pi}_{T_k}) - \mathbb{E}\left[u(T_{k+1},X^{T_k,\hat{X}^{\pi}_{T_k},\mu}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\bar{\Theta}^{\mu,\mu}_{k+1,k}(\hat{X}^{\pi}_{T_k}))\right]\right| \\
&\quad \leq C \sup_{s\in[T_k,T_{k+1}]} ||u(s,\cdot)||_{4,\infty}\Delta^2_{T_{k+1}} \\
&\left|\Delta_{T_{k+1}} v(T_k,\hat{X}^{\pi}_{T_k}) - \mathbb{E}\left[u\left(T_{k+1},X^{T_k,\hat{X}^{\pi}_{T_k},\mu}_{T_{k+1}}\right)\Delta B_{T_{k+1}}\right]\right| \\
&\quad \leq C \sup_{s\in[T_k,T_{k+1}]} ||u(s,\cdot)||_{3,\infty}\Delta^2_{T_{k+1}}
\end{aligned}$$

*Proof.* The proof of the first estimate follows from a second order Itô-Taylor expansion. Applying Lemma 2.5.1 with $n=2$ to $u$ and taking the expectation leads to :

$$\mathbb{E}\left[u(T_{k+1},X^{T_k,y,\mu}_{T_{k+1}})\right] = u(T_k,y) + \Delta_{T_k}\mathcal{V}_{(0)}u(T_k,y) + \sum_{\beta\in\partial\mathcal{A}_2}\mathbb{E}\left(I^{T_k,T_{k+1}}_{\beta}[\mathcal{V}_{\beta}u(\cdot,X^{T_k,y,\mu}_{\cdot})]\right) \tag{2.6.17}$$

and applying again Lemma 2.5.1 with $n=1$ to $\mathcal{V}_{(0)}u$ and taking the expectation gives :

$$\mathbb{E}\left[\mathcal{V}_{(0)}u(T_{k+1},X^{T_k,y,\mu}_{T_{k+1}})\right] = \mathcal{V}_{(0)}u(T_k,y) + \sum_{\beta\in\partial\mathcal{A}_1}\mathbb{E}\left(I^{T_k,T_{k+1}}_{\beta}[\mathcal{V}_{\beta}\mathcal{V}_{(0)}u(\cdot,X^{T_k,y,\mu}_{\cdot})]\right). \tag{2.6.18}$$

Now, note that since $u$ is the solution of PDE (2.0.9) we have $f=-\mathcal{V}_{(0)}u$. So that, by combining (2.6.17), (2.6.18) and estimate of Lemma 2.5.2, we obtain

$$\begin{aligned}
&\left|u(T_k,\hat{X}^{\pi}_{T_k}) - \mathbb{E}\left[u(T_{k+1},X^{T_k,\hat{X}^{\pi}_{T_k},\mu}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\bar{\Theta}^{\mu,\mu}_{k+1,k}(\hat{X}^{\pi}_{T_k}))\right]\right| \\
&\leq 2\sum_{\beta\in\partial\mathcal{A}_2}\mathbb{E}\left(I^{T_k,T_{k+1}}_{\beta}[\mathcal{V}_{\beta}u(\cdot,X^{T_k,y,\mu}_{\cdot})]\right) \\
&\leq C(T,V_{0:d},f)\sup_{s\in[T_k,T_{k+1}]}||u(s,\cdot)||_{4,\infty}\Delta^2_{T_{k+1}}.
\end{aligned}$$

This concludes the proof of the first estimate. The proof of the second estimate is similar. We first apply an Itô-Taylor expansion of Lemma 2.5.1 with $n=1$ on $u$. Then, by noticing that $\Delta B_{T_{k+1}} = (I^{T_k,T_{k+1}}_{(1)},\dots,I^{T_k,T_{k+1}}_{(d)})^T$ and by multiplying by $I^{T_k,T_{k+1}}_{(j)}$ the previous expansion of $u$, and taking the expectation gives, thanks to Itô's Formula

$$\mathbb{E}\left[u(T_{k+1},X^{T_k,y,\mu}_{T_{k+1}})I^{T_k,T_{k+1}}_{(j)}\right] = \mathcal{V}_{(j)}u(T_k,y)\Delta_{T_{k+1}} + \sum_{\beta\in\partial\mathcal{A}_1}\mathbb{E}\left(I^{T_k,T_{k+1}}_{\beta}[\mathcal{V}_{\beta}u(\cdot,X^{T_k,y,\mu}_{\cdot})]I^{T_k,T_{k+1}}_{(j)}\right),$$

for $j=1,\dots,d$ and where the first term in the right hand side is the bracket between the stochastic integrals. The last term is controlled by using Lemma 2.5.2. Recalling that the $j$-th component of the function $v$ is given by $\mathcal{V}_{(j)}u$ and reordering the terms we obtain the second inequality. □

We now turn to bound the cubature like error terms (2.6.12) and (2.6.15). This is summarized by :

**Claim 2.6.5.** *There exist two constants $C$, depending on $d$, $q$, $T$, $m$, and the regularity of $V_{0:d}$ and $\varphi_{0:d}$ (and not on $k$), and $C'$, depending in addition on the regularity of $f$, such that :*

$$
\begin{aligned}
&\Bigg|\mathbb{E}\left[u(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\mu}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\bar{\Theta}^{\mu,\mu}_{k+1,k}(\hat{X}^{\pi}_{T_k}))\right] \\
&\quad - \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\left[u(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\bar{\Theta}^{\hat{\mu},\mu}_{k+1,k}(\hat{X}^{\pi}_{T_k}))\right]\Bigg| \\
&\quad \leq C\Bigg(||u(T_{k+1},\cdot)||_{2,\infty}\left[\Delta^{q+1}_{T_{k+1}} + \Delta_{T_{k+1}} N^{-[(m-1)\wedge 2q]/2}\right] + ||u(T_{k+1},\cdot)||_{m+1,\infty}\Delta^{(m+1)/2}_{T_{k+1}} \\
&\quad + ||u(T_{k+1},\cdot)||_{m+2,\infty}\Delta^{(m+2)/2}_{T_{k+1}}\Bigg) + C'\Bigg(M_u(2,T_{k+1})\left[\Delta^{q+1}_{T_{k+1}} + \Delta_{T_{k+1}} N^{-[(m-1)\wedge 2q]/2}\right] \\
&\quad + M_u(m+1,T_{k+1})\Delta^{(m+1)/2}_{T_{k+1}} + M_u(m+2,T_{k+1})\Delta^{(m+2)/2}_{T_{k+1}}\Bigg) \\
&\Bigg|\mathbb{E}\left[u\left(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\mu}_{T_{k+1}}\right)\Delta B_{T_{k+1}}\right] - \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\left[u\left(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}}\right)\Delta B_{T_{k+1}}\right]\Bigg| \\
&\quad \leq C\Bigg(||u(T_{k+1},\cdot)||_{3,\infty}\left[\Delta^{q+1}_{T_{k+1}} + \Delta^{2}_{T_{k+1}} N^{-[(m-1)\wedge 2q]/2}\right] + ||u(T_{k+1},\cdot)||_{m,\infty}\Delta^{(m+1)/2}_{T_{k+1}} \\
&\quad + ||u(T_{k+1},\cdot)||_{m+1,\infty}\Delta^{(m+2)/2}_{T_{k+1}}\Bigg)
\end{aligned}
$$

*Proof.* Note that the $r^{\text{th}}$ derivative of the function $y \mapsto f(\cdot, y, u(\cdot, y), \cdot, \cdot)$ is bounded by $C'M_u(r, \cdot)$ defined by (2.6.10). Then, the proof of the first assertion follows from (2.7.14) in Lemma 2.7.5 applied to $u$ and $f$ and the second assertion from (2.7.15) in Lemma 2.7.5 applied to $u$. □

Finally, an estimate on the propagation error (2.6.13) is given by :

**Claim 2.6.6.** *There exists a constant $C$ depending on $d$, $q$, $T$, $m$, and the regularity of $V_{0:d}$ and $\varphi_{0:d}$ (and not on $k$) such that :*

$$
\begin{aligned}
&\Bigg|\mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\Bigg[u(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}}) - \hat{u}^1(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}}) \\
&\quad + \Delta_{T_{k+1}}\left(f(\bar{\Theta}^{\hat{\mu},\mu}_{k+1,k}(\hat{X}^{\pi}_{T_k})) - f(\hat{\Theta}^{\pi,1}_{k+1,k})\right)\Bigg]\Bigg| \\
&\quad \leq (1 + C\Delta_{T_{k+1}})\mathcal{E}^1_u(k+1) + C\Bigg(||u(T_{k+1},\cdot)||_{m+2,\infty}\Delta_{T_{k+1}} N^{-[(m-1)/2]\wedge q} \\
&\quad + ||u(T_{k+1},\cdot)||_{3,\infty}\left[\Delta^2_{T_{k+1}} + \Delta^{q+1}_{T_{k+1}} + \Delta^2_{T_{k+1}} N^{-[(m-1)\wedge 2q]/2}\right] \\
&\quad + ||u(T_{k+1},\cdot)||_{m+1,\infty}\Delta^{(m+1)/2}_{T_{k+1}} + ||u(T_{k+1},\cdot)||_{m+2,\infty}\Delta^{(m+2)/2}_{T_{k+1}}\Bigg)
\end{aligned}
$$

*Proof.* Let us start by expanding the $f$ term. We get from the mean value theorem that there exist three random variable $\Psi_1, \Psi_2, \Psi_3$, respectively bounded by $\|\partial_{y'} f\|_\infty$, $\|\partial_z f\|_\infty$

and $\|\partial_w f\|_\infty$ almost surely, depending on each argument of $\hat{\Theta}^{\pi,1}_{k+1,k}$ and $\bar{\Theta}^{\hat{\mu},\mu}_{k+1,k}$ such that :

$$
\begin{aligned}
&\Delta_{T_{k+1}} \left( f(\bar{\Theta}^{\hat{\mu},\mu}_{k+1,k}(\hat{X}^\pi_{T_k})) - f(\hat{\Theta}^{\pi,1}_{k+1,k}) \right) \\
&= \Delta_{T_{k+1}} \Psi_1 \left( u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) - \hat{u}^1(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) \right) \\
&\quad + \Delta_{T_{k+1}} \Psi_2 \left( v(T_k, \hat{X}^\pi_{T_k}) - \hat{v}^1(T_k, \hat{X}^\pi_{T_k}) \right) \\
&\qquad + \Delta_{T_{k+1}} \Psi_3 \left( \langle \mu_{T_{k+1}}, \varphi_f[\cdot, u(T_{k+1}, \cdot)] \rangle - \langle \hat{\mu}_{T_{k+1}}, \varphi_f[\cdot, \hat{u}^1(T_{k+1}, \cdot)] \rangle \right).
\end{aligned}
$$

Now, we can use the error expansion (2.6.14), (2.6.15) and (2.6.16) of

$$
\Delta_{T_{k+1}} \left( v(T_k, \hat{X}^\pi_{T_k}) - \hat{v}^1(T_k, \hat{X}^\pi_{T_k}) \right),
$$

together with the second assertion of Claims 2.6.4 and 2.6.5 to get

$$
\begin{aligned}
&\Bigg| \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \Bigg[ u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) - \hat{u}^1(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) \qquad (2.6.19)\\
&\qquad + \Delta_{T_{k+1}} \left( f(\bar{\Theta}^{\hat{\mu},\mu}_{k+1,k}(\hat{X}^\pi_{T_k})) - f(\hat{\Theta}^{\pi,1}_{k+1,k}) \right) \Bigg] \Bigg| \\
&\leq \left| \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \left[ \left( u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) - \hat{u}^1(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) \right) (1 + \Psi_1 \Delta_{T_{k+1}} + \Psi_2 \Delta B_{T_{k+1}}) \right] \right| \\
&\quad + C ||u(T_{k+1}, \cdot)||_{3,\infty} \left[ \Delta^2_{T_{k+1}} + \Delta^{q+1}_{T_{k+1}} + \Delta^2_{T_{k+1}} N^{-[(m-1)\wedge 2q]/2} \right] \\
&\quad + C ||u(T_{k+1}, \cdot)||_{m+1,\infty} \Delta^{(m+1)/2}_{T_{k+1}} + C ||u(T_{k+1}, \cdot)||_{m+2,\infty} \Delta^{(m+2)/2}_{T_{k+1}} \\
&\quad + \Delta_{T_{k+1}} \|\partial_w f\|_\infty \left| \langle \mu - \hat{\mu}_{T_{k+1}}, \varphi_f[\cdot, u(T_{k+1}, \cdot)] \rangle \right| \\
&\quad + \Delta_{T_{k+1}} \|\partial_w f\|_\infty \left| \langle \hat{\mu}_{T_{k+1}}, \varphi_f[\cdot, u(T_{k+1}, \cdot)] - \varphi_f[\cdot, \hat{u}^1(T_{k+1}, \cdot)] \rangle \right|. \qquad (2.6.20)
\end{aligned}
$$

Note that from the forward result (2.2.2) in Theorem 2.2.1, we have

$$
\Delta_{T_{k+1}} \|\partial_w f\|_\infty \left| \langle \mu - \hat{\mu}_{T_{k+1}}, \varphi_f[\cdot, u(T_{k+1}, \cdot)] \rangle \right| \leq C(||u(T_{k+1}, \cdot)||_{m+2,\infty}) \Delta_{T_{k+1}} N^{-((m-1)\wedge 2q)/2}, \tag{2.6.21}
$$

while using the regularity of $\varphi_f$ and the definition of $\hat{\mu}$ gives,

$$
\Delta_{T_{k+1}} \|\partial_w f\|_\infty \left| \langle \hat{\mu}_{T_{k+1}}, \varphi_f[\cdot, u(T_{k+1}, \cdot)] - \varphi_f[\cdot, \hat{u}^1(T_{k+1}, \cdot)] \rangle \right| \leq C' \Delta_{T_{k+1}} \mathcal{E}^1_u(k+1). \tag{2.6.22}
$$

The Claim follows by applying the Cauchy-Schwartz inequality on the first term in the right hand side of (2.6.20) and plugging (2.6.21) and (2.6.22) in (2.6.20). $\square$

We can now analyze the local error at step $k$. By plugging the estimates from Claims 2.6.4, 2.6.5 and 2.6.6 in the expansion (2.6.11), (2.6.12) and (2.6.13) of $u - \hat{u}$ we obtain that :

$$
\mathcal{E}^1_u(k) \leq \left(1 + C\Delta_{T_{k+1}}\right) \mathcal{E}^1_u(k+1) + \bar{\epsilon}(k+1), \tag{2.6.23}
$$

with

$$\begin{aligned}
\bar{\epsilon}(k+1) &= C\Bigg(\sup_{s\in[T_k;T_{k+1}]}||u(s,\cdot)||_{2,\infty}\Delta^4_{T_{k+1}} + ||u(T_{k+1},\cdot)||_{2,\infty}\Delta_{T_{k+1}}N^{-[(m-1)\wedge 2q]/2} \\
&\quad +||u(T_{k+1},\cdot)||_{m+1,\infty}\Delta^{(m+1)/2}_{T_{k+1}} + ||u(T_{k+1},\cdot)||_{m+2,\infty}\Delta^{(m+2)/2}_{T_{k+1}} \\
&\quad +||u(T_{k+1},\cdot)||_{3,\infty}\left[\Delta^2_{T_{k+1}} + \Delta^{q+1}_{T_{k+1}} + \Delta^2_{T_{k+1}}N^{-[(m-1)\wedge 2q]/2}\right]\Bigg) \qquad (2.6.24)\\
&\quad +C'\Bigg(M_u(m+1,T_{k+1})\Delta^{(m+1)/2}_{T_{k+1}} + M_u(m+2,T_{k+1})\Delta^{(m+2)/2}_{T_{k+1}} \\
&\quad +M_u(2,T_{k+1})\left[\Delta^{q+1}_{T_{k+1}} + \Delta_{T_{k+1}}N^{-[(m-1)\wedge 2q]/2}\right]\Bigg).
\end{aligned}$$

Under **(SB)** , we have from Lemma 2.7.2 that for all $n\in\mathbb{N}^*$ there exists a constant $K$, depending on the regularity of $V_{0:d}$ and $\phi$, such that for all $k\in\{0,\cdots,N-1\}$,

$$M_u(n,T_{k+1}) + \sup_{s\in[T_k,T_{k+1}]}||u(s,\cdot)||_{n,\infty} \leq K.$$

Therefore, Gronwall's Lemma applied to (2.6.23) and the definition of $\Delta_{T_k}$ implies

$$\mathcal{E}^1_u \leq CN^{-1}. \qquad (2.6.25)$$

Moreover, Claims 2.6.4 and 2.6.5 and expansion (2.6.14), (2.6.15) and (2.6.16) show that :

$$\begin{aligned}
\Delta_{T_{k+1}}\mathcal{E}^1_v(k+1) \leq C\Bigg(&\sup_{s\in[T_k,T_{k+1}]}||u(s,\cdot)||\Delta^2_{T_{k+1}} \\
&+||u(T_{k+1},\cdot)||_{3,\infty}\left[\Delta^{q+1}_{T_{k+1}} + \Delta^2_{T_{k+1}}N^{-[(m-1)\wedge 2q]/2}\right] \\
&+||u(T_{k+1},\cdot)||_{m+1,\infty}\Delta^{(m+1)/2}_{T_{k+1}} + C||u(T_{k+1},\cdot)||_{m+2,\infty}\Delta^{(m+2)/2}_{T_{k+1}}\Bigg) \\
&+\Delta^{1/2}_{T_{k+1}}\mathcal{E}^1_u(k+1),
\end{aligned}$$

which together with (2.6.25) imply

$$\Delta^{1/2}_{T_k}\mathcal{E}^1_v \leq CN^{-1},$$

and the result holds.

*(2) Proof of the order of convergence for the second order algorithm (Algorithm 3).*

Let $k\in\{0,\ldots,N-2\}$. We can expand as before the errors on the $u$ and $v$ approximations as :

$$
\begin{aligned}
&u(T_k, \hat{X}^\pi_{T_k}) - \hat{u}^2(T_k, \hat{X}^\pi_{T_k}) \\
&= \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \left[ u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) - \hat{u}^2(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) \right] \\
&\qquad + \frac{1}{2}\Delta_{T_{k+1}} \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \left[ f(\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}})) - f(\hat{\Theta}^{\pi,2}_{k+1}) \right] \qquad (2.6.26) \\
&\quad + \frac{1}{2}\Delta_{T_{k+1}} [f(\Theta^\mu_k(\hat{X}^\pi_{T_k})) - f(\tilde{\Theta}^\pi_k)] \qquad (2.6.27) \\
&\quad + \epsilon^s_{\hat{u}^2,k}(\pi) + \epsilon^c_{\hat{u}^2,k}(\pi),
\end{aligned}
$$

where

$$
\begin{aligned}
\epsilon^s_{\hat{u}^2,k}(\pi) &= u(T_k, \hat{X}^\pi_{T_k}) - \mathbb{E}\Big[ u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}}) \\
&\qquad + \frac{\Delta_{T_{k+1}}}{2} \left( f(\Theta^\mu_k(\hat{X}^\pi_{T_k})) + f(\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}})) \right) \Big] \qquad (2.6.28) \\
\epsilon^c_{\hat{u}^2,k}(\pi) &= \mathbb{E}\left[ u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}}) \right] - \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \left[ u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) \right] \\
&\quad + \frac{\Delta_{T_{k+1}}}{2} \left( \mathbb{E}\left[ f(\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}})) \right] - \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \left[ f(\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}})) \right] \right). \qquad (2.6.29)
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&v(T_k, \hat{X}^\pi_{T_k}) - \hat{v}^2(T_k, \hat{X}^\pi_{T_k}) \\
&= \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \left( \left[ u\left(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}\right) - \hat{u}^2\left(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}\right) \right] \zeta_{k+1} \right) \\
&\quad + \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \left[ \left( f[\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}})] - f(\hat{\Theta}^{\pi,2}_{k+1}) \right) \Delta_{T_{k+1}} \zeta_{k+1} \right] \qquad (2.6.30) \\
&\quad + \epsilon^s_{\hat{v}^2,k}(\pi) + \epsilon^c_{\hat{v}^2,k}(\pi),
\end{aligned}
$$

with

$$
\epsilon^s_{\hat{v}^2,k}(\pi) = v(T_k, \hat{X}^\pi_{T_k}) - \mathbb{E}\left( \left[ u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}}) + \Delta_{T_{k+1}} f[\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}})] \right] \zeta_{k+1} \right) \qquad (2.6.31)
$$

$$
\begin{aligned}
\epsilon^c_{\hat{v}^2,k}(\pi) &= \mathbb{E}\left[ u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}}) \zeta_{k+1} \right] - \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \left[ u(T_{k+1}, X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}}) \zeta_{k+1} \right] \\
&\quad + \mathbb{E}\left[ f[\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}})] \Delta_{T_{k+1}} \zeta_{k+1} \right] - \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}} \left[ f[\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}})] \Delta_{T_{k+1}} \zeta_{k+1} \right]. \qquad (2.6.32)
\end{aligned}
$$

We identify, as for the first order expansion, some error terms corresponding to the scheme error (2.6.28) and (2.6.31), generalized cubature errors (2.6.29) and (2.6.32) and propagation errors (2.6.26), (2.6.30). Some important changes are clear from the expansion : we have in addition a prediction error term (2.6.27) reflecting the fact that we perform a new intermediate step, and we have some $f$ term in (2.6.30), adding to the propagation error.

The proof for the second order approximation is then similar to its first order equivalent, but we will have to consider the mentioned additional terms. In particular, the fact that the second order approximation $\hat{v}^2(\hat{X}^\pi_{T_k})$ includes the term $f(\hat{\Theta}^{\pi,2}_{T_{k+1}})$, adds an additional coupling effect. With this in mind, and in order to simplify the analysis, we introduce the following quantity

$$\mathcal{E}_f(k) = \max_{\pi\in\mathcal{S}(k)} |f(\Theta^\mu_k(\hat{X}^\pi_{T_k})) - f(\hat{\Theta}^{\pi,2}_k)|$$

and we will analyze the dynamics of the sum of errors at step $k$, $\mathcal{E}^2_u(k) + \Delta_{T_k}\mathcal{E}_f(k)$.

To this aim, we will bound separately the scheme and cubature errors. Each bound is summarized in a Claim (respectively Claims 2.6.7, 2.6.8 below). Then, we will conclude with a Gronwall argument.

The scheme error terms (2.6.28) and (2.6.31) and the generalized cubature errors (2.6.29) and (2.6.32) are treated similarly as in the first order scheme. We show this in Claims 2.6.7 and 2.6.8.

**Claim 2.6.7.** *There exists a constant $C$, depending on the regularity of $V_{0:d}$ and $f$ (and not on $k$) such that :*

$$\begin{aligned}
\left|\epsilon^s_{\hat{u}^2,k}(\pi)\right| &\leq C \sup_{s\in[T_k,T_{k+1}]} ||u(s,\cdot)||_{6,\infty}\Delta^3_{T_{k+1}}\\
\left|\epsilon^s_{\hat{v}^2,k}(\pi)\right| &\leq C \sup_{s\in[T_k,T_{k+1}]} ||u(s,\cdot)||_{5,\infty}\Delta^2_{T_{k+1}}.
\end{aligned}$$

*where $\epsilon^s_{\hat{u}^2,k}(\pi), \epsilon^s_{\hat{v}^2,k}(\pi)$ are defined in (2.6.28) and (2.6.31).*

*Proof.* The proof follows in the same way as the one of Claim 2.6.4, by performing a Taylor expansion to one additional order. The choice of $\zeta_{k+1}$ is the one needed to match the lower order terms (recall that the $i^{th}$ component of $\zeta_k$ is expressed as $\zeta^i_k = 4\Delta^{-1}_{T_k} I^{T_k,T_{k+1}}_{(i)} - 6\Delta^{-2}_{T_k} I^{T_k,T_{k+1}}_{(0,i)}$).

Applying Lemma 2.5.1 with $n = 4$ to $u$ and with $n = 2$ to $\mathcal{V}_{(0)}u$ implies, after taking the expectation, that

$$\begin{aligned}
\mathbb{E}\left[u(T_{k+1}, X^{T_k,y,\mu}_{T_{k+1}})\right] &= u(T_k,y) + \Delta_{T_k}\mathcal{V}_{(0)}u(T_k,y) + \frac{1}{2}\mathcal{V}_{(0,0)}u(T_k,y)\Delta^2_{T_{k+1}}\\
&\quad + \sum_{\beta\in\partial\mathcal{A}_4} \mathbb{E}\left(I^{T_k,T_{k+1}}_\beta[\mathcal{V}_\beta u(.,X^{T_k,y,\mu}_.)]\right)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\left[\mathcal{V}_{(0)}u(T_{k+1}, X^{T_k,y,\mu}_{T_{k+1}})\right] &= \mathcal{V}_{(0)}u(T_k,y) + \mathcal{V}_{(0,0)}u(T_k,y)\Delta_{T_{k+1}}\\
&\quad + \sum_{\beta\in\partial\mathcal{A}_2} \mathbb{E}\left(I^{T_k,T_{k+1}}_\beta[\mathcal{V}_{(\beta*0)}u(.,X^{T_k,y,\mu}_.)]\right).
\end{aligned}$$

Then, the estimate of Lemma 2.5.2 gives

$$\begin{aligned}
&\mathbb{E}\left[u(T_{k+1}, X^{T_k,y,\mu}_{T_{k+1}}) - \frac{\Delta_{T_k}}{2}\left(\mathcal{V}_{(0)}u(T_k,y) + \mathcal{V}_{(0)}u(T_{k+1}, X^{T_k,y,\mu}_{T_{k+1}})\right)\right] - u(T_k,y)\\
&= \sum_{\beta\in\partial\mathcal{A}_4} \mathbb{E}\left(I^{T_k,T_{k+1}}_\beta[\mathcal{V}_{(\beta)}u(.,X^{T_k,y,\mu}_.)]\right) - \frac{\Delta_{T_k}}{2}\sum_{\beta\in\partial\mathcal{A}_2} \mathbb{E}\left(I^{T_k,T_{k+1}}_\beta[\mathcal{V}_{(\beta*0)}u(.,X^{T_k,y,\mu}_.)]\right)\\
&\leq C(T,V_{0:d},f)\Delta^3_{T_{k+1}} \sup_{s\in[T_k,T_{k+1}]} ||u(s,.)||_{6,\infty},
\end{aligned}$$

from where we deduce the first inequality.

Similarly, by using Lemma 2.5.1 with $n=3$ on $u$ and with $n=1$ on $\mathcal{V}_{(0)}u$ and taking the expectation, using the fact that $\zeta_k^i = 4\Delta_{T_k}^{-1} I_{(i)}^{T_k,T_{k+1}} - 6\Delta_{T_k}^{-2} I_{(0,i)}^{T_k,T_{k+1}}$, it follows

$$
\begin{aligned}
&\mathbb{E}\left[\left[u(T_{k+1}, X_{T_{k+1}}^{T_k,y,\mu}) - \Delta_{T_{k+1}}\mathcal{V}_{(0)}u(T_{k+1}, X_{T_{k+1}}^{T_k,y,\mu})\right]\Delta_{T_{k+1}}\zeta_k^j\right] \\
&= \mathbb{E}\Bigg[\Big[\mathcal{V}_{(j)}u(T_k,y)I_{(j)}^{T_k,T_{k+1}} + \mathcal{V}_{(0,j)}u(T_k,y)I_{(0,j)}^{T_k,T_{k+1}} + \mathcal{V}_{(j,0)}u(T_k,y)I_{(j,0)}^{T_k,T_{k+1}} \\
&\qquad - \Delta_{T_{k+1}}\mathcal{V}_{(j,0)}u(T_k,y)I_{(j)}^{T_k,T_{k+1}}\Big]\left(4I_{(j)}^{T_k,T_{k+1}} - 6\frac{I_{(0,j)}^{T_k,T_{k+1}}}{\Delta_{T_{k+1}}}\right)\Bigg] + \mathcal{R}(k,j) \\
&= \Delta_{T_{k+1}}\mathcal{V}_{(j)}u(T_k,y) + \mathcal{R}(k,j)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{R}(k,j) = \Delta_{T_{k+1}} &\sum_{\beta\in\partial\mathcal{A}_3} \mathbb{E}\left(I_\beta^{T_k,T_{k+1}}[\mathcal{V}_{(\beta)}u(.,X_.^{T_k,y,\mu})]\zeta_k^j\right) \\
&- \Delta_{T_{k+1}}^2 \sum_{\beta\in\partial\mathcal{A}_1} \mathbb{E}\left(I_\beta^{T_k,T_{k+1}}[\mathcal{V}_{(\beta*0)}u(.,X_.^{T_k,y,\mu})]\zeta_k^j\right).
\end{aligned}
$$

Using Lemma 2.5.2 we bound the residual term $\mathcal{R}(k,j)$ and obtain

$$
\begin{aligned}
&\Delta_{T_{k+1}} \sum_{\beta\in\partial\mathcal{A}_3} \mathbb{E}\left(I_\beta^{T_k,T_{k+1}}[\mathcal{V}_{(\beta)}u(.,X_.^{T_k,y,\mu})]\zeta_k^j\right) - \Delta_{T_{k+1}}^2 \sum_{\beta\in\partial\mathcal{A}_1} \mathbb{E}\left(I_\beta^{T_k,T_{k+1}}[\mathcal{V}_{(\beta*0)}u(.,X_.^{T_k,y,\mu})]\zeta_k^j\right) \\
&\quad \le \Delta_{T_k}^3 \sup_{s\in[T_k,T_{k+1}]} ||u(s,.)||_{5,\infty} + \Delta_{T_k}^{7/2} \sup_{s\in[T_k,T_{k+1}]} ||u(s,.)||_{5,\infty},
\end{aligned}
$$

recalling that $v(T_k,y) = \mathcal{V}_{(j)}u(T_k,y)$ we deduce the second inequality. □

**Claim 2.6.8.** *There exist two constants $C$, depending on $d$, $q$, $T$, $m$, the regularity of $V_{0:d}$ and $\varphi_{0:d}$ (and not on $k$), and $C'$, depending in addition on the regularity of $f$, such that :*

$$
\begin{aligned}
\left|\epsilon^c_{\hat{u}^2,k}(\pi)\right| \le C\Bigg(&||u(T_{k+1},\cdot)||_{2,\infty}\left[\Delta_{T_{k+1}}^{q+1} + \Delta_{T_{k+1}}N^{-[(m-1)\wedge 2q]/2}\right] + ||u(T_{k+1},\cdot)||_{m+1,\infty}\Delta_{T_{k+1}}^{(m+1)/2} \\
&+ ||u(T_{k+1},\cdot)||_{m+2,\infty}\Delta_{T_{k+1}}^{(m+2)/2}\Bigg) \\
&+ C'\Bigg(M_u(3,T_{k+1})\left[\Delta_{T_{k+1}}^{q+2} + \Delta_{T_{k+1}}^2 N^{-[(m-1)\wedge 2q]/2}\right] + M_u(m+2,T_{k+1})\Delta_{T_{k+1}}^{(m+3)/2} \\
&+ M_u(m+3,T_{k+1})\Delta_{T_{k+1}}^{(m+4)/2}\Bigg),
\end{aligned}
$$

$$
\begin{aligned}
\left|\epsilon^c_{\hat{v}^2,k}(\pi)\right| \le C\Bigg(&||u(T_{k+1},\cdot)||_{3,\infty}\left[\Delta_{T_{k+1}}^{q} + \Delta_{T_{k+1}}N^{-[(m-1)\wedge 2q]/2}\right] + \sum_{i=m-2}^{m+1}||u(T_{k+1},\cdot)||_{i,\infty}\Delta_{T_{k+1}}^{(i-1)/2}\Bigg) \\
&+ C'\Bigg(M_u(4,T_{k+1})\left[\Delta_{T_{k+1}}^{q+1} + \Delta_{T_{k+1}}^2 N^{-[(m-1)\wedge 2q]/2}\right] + \sum_{i=m-1}^{m+2} M_u(i,T_{k+1})\Delta_{T_{k+1}}^{i/2}\Bigg).
\end{aligned}
$$

*where $\epsilon^c_{\hat{u}^2,k}(\pi), \epsilon^c_{\hat{v}^2,k}(\pi)$ are defined in (2.6.29) and (2.6.32).*

**Remark.** *Although the rates of convergence have a leading term of order $\Delta_{T_{k+1}}^{(m-1)}$ that is worst than the one in the first order scheme result, (Claim 2.6.5 ), here we assume that $m$ is bigger, and thus they are suitable for a second order scheme.*

*Proof.* Note first that the $r^{\text{th}}$ derivative of the function $y \mapsto f(\cdot, y, u(\cdot, y), \mathcal{V}u(\cdot, y), \cdot)$ is bounded by $M_u(r+1, \cdot)$ defined by (2.6.10). This estimate goes up to $r+1$ and not just $r$ as in Claim 2.6.5, because the differentiation of $y \in \mathbb{R}^d \mapsto f(\Theta^{\mu}_{k+1}(y))$ involves the additional dependence on $\mathcal{V}u$.

Then, the first assertion follows from applying (2.7.14) in Lemma 2.7.5 to $u$ and $f(\Theta^{\mu}_{k+1})$. For the second assertion, recall that the $i^{th}$ component of $\zeta_k$ is expressed as $\zeta^i_k = 4\Delta^{-1}_{T_k} I^{T_k,T_{k+1}}_{(i)} - 6\Delta^{-2}_{T_k} I^{T_k,T_{k+1}}_{(0,i)}$. Then, applying (2.7.15) and (2.7.16) with $n = m$ in Lemma 2.7.5 to $u$ and $f(\Theta^{\mu}_{k+1})$, we conclude on the second assertion. $\square$

It will be handy to have an expansion on the prediction error, by recalling that $\tilde{u}$ is essentially an application of the first order scheme, it follows that

$$\begin{aligned}(u - \tilde{u})(T_k, \hat{X}^{\pi}_{T_k}) &= \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\Big[(u - \hat{u}^2)(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}}) \\ &\quad + \Delta_{T_{k+1}}\Big(f(\Theta^{\hat{\mu}}_{k+1}(X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}}) - f(\hat{\Theta}^{\pi,2}_{k+1})\Big)\Big] \\ &\quad + \epsilon^s_{\tilde{u},k}(\pi) + \epsilon^c_{\tilde{u},k}(\pi)\end{aligned} \tag{2.6.33}$$

where

$$\begin{aligned}\epsilon^c_{\tilde{u},k}(\pi) &:= \mathbb{E}\Big[u(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\mu}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\Theta^{\mu}_{k+1}(X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}})\Big] \\ &\quad - \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\Big[u(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}}) + \Delta_{T_{k+1}} f(\Theta^{\hat{\mu}}_{k+1}(X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}})\Big]\end{aligned} \tag{2.6.34}$$

$$\epsilon^s_{\tilde{u},k}(\pi) := u(T_k, \hat{X}^{\pi}_{T_k}) - \mathbb{E}\Big[u(T_{k+1}, X^{T_k,\hat{X}^{\pi}_{T_k},\mu}_{T_{k+1}}) + \Delta_{T_{k+1}} f[\Theta^{\mu}_{k+1}(X^{T_k,\hat{X}^{\pi}_{T_k},\hat{\mu}}_{T_{k+1}})]\Big]. \tag{2.6.35}$$

Note that $\epsilon^s_{\tilde{u},k}(\pi)$ and $\epsilon^c_{\tilde{u},k}(\pi)$ may be bounded respectively as in Claims (2.6.4) and (2.6.5) . The fact that we are using here $\Theta^{\mu}_{k+1}$ instead of $\bar{\Theta}^{\mu,\mu}_{k+1,k}$ in those claims, is not really problematic. For the scheme error in Claim (2.6.4), the difference can be controlled by an additional application of Ito's theorem, but we skip the details. For the cubature error in Claim (2.6.5) this difference plays no role at all.

Let us now focus on the errors when approaching the driver. Using the mean value theorem, we know that there exist $\Psi_1, \Psi_2, \Psi_3$ respectively bounded by $\|\partial_{y'} f\|_\infty$, $\|\partial_z f\|_\infty$ and $\|\partial_w f\|_\infty$, such that

$$f(\Theta^{\mu}_k(\hat{X}^{\pi}_{T_k})) - f(\hat{\Theta}^{\pi,2}_k) = \Psi_1\big[u(T_k, \hat{X}^{\pi}_{T_k}) - \hat{u}^2(T_k, \hat{X}^{\pi}_{T_k})\big] + \Psi_2\big[v(T_k, \hat{X}^{\pi}_{T_k}) - \hat{v}^2(T_k, \hat{X}^{\pi}_{T_k})\big] \tag{2.6.36}$$

$$+ \Psi_3\left(\langle \hat{\mu}_{T_k}, \varphi_f[\cdot, u(T_k, \cdot)]\rangle - \langle \hat{\mu}_{T_k}, \varphi_f[\cdot, \hat{u}^2(T_k, \cdot), \cdot)]\rangle\right) \tag{2.6.37}$$

$$+ \Psi_3\left(\langle \mu_{T_k}, \varphi_f[\cdot, u(T_k, \cdot)]\rangle - \langle \hat{\mu}_{T_k}, \varphi_f[\cdot, u(T_k, \cdot), \cdot)]\rangle\right), \tag{2.6.38}$$

Similarly, there exist random variables and $\Psi'_1, \Psi'_2, \Psi'_3$ respectively bounded by $\|\partial_{y'} f\|_\infty$, $\|\partial_z f\|_\infty$ and $\|\partial_w f\|_\infty$, such that

$$f(\Theta^{\mu}_k(\hat{X}^{\pi}_{T_k})) - f(\tilde{\Theta}^{\pi}_k) = \Psi'_1\big[u(T_k, \hat{X}^{\pi}_{T_k}) - \tilde{u}(T_k, \hat{X}^{\pi}_{T_k})\big] + \Psi'_2\big[v(T_k, \hat{X}^{\pi}_{T_k}) - \hat{v}^2(T_k, \hat{X}^{\pi}_{T_k})\big] \tag{2.6.39}$$

$$+ \Psi'_3\left(\langle \hat{\mu}_{T_k}, \varphi_f[\cdot, u(T_k, \cdot)]\rangle - \langle \hat{\mu}_{T_k}, \varphi_f[\cdot, \tilde{u}(T_k, \cdot), \cdot)]\rangle\right) \tag{2.6.40}$$

$$+ \Psi'_3\left(\langle \mu_{T_k}, \varphi_f[\cdot, u(T_k, \cdot)]\rangle - \langle \hat{\mu}_{T_k}, \varphi_f[\cdot, u(T_k, \cdot), \cdot)]\rangle\right).. \tag{2.6.41}$$

Then, using the error development for $(u-\hat{u}^2)(T_k,\hat{X}^\pi_{T_k})$ given in (2.6.28),(2.6.29), (2.6.26),(2.6.27) the error and the ones for $f(\Theta^\mu_k(\hat{X}^\pi_{T_k}))-f(\hat{\Theta}^{\pi,2}_k)$ in (2.6.36),(2.6.37) and (2.6.38) and $f(\Theta^\mu_k(\hat{X}^\pi_{T_k}))-f(\tilde{\Theta}^\pi_k)$ in (2.6.39), (2.6.40) and (2.6.41) , it follows

$$
\begin{aligned}
&(u-\hat{u}^2)(T_k,\hat{X}^\pi_{T_k})+\frac{1}{2}\Delta_{T_k}\left(f(\Theta^\mu_k(\hat{X}^\pi_{T_k}))-f(\hat{\Theta}^{\pi,2}_k)\right)\\
&\qquad=(1+\Delta_{T_k}\Psi_1)\mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\Big[u(T_{k+1},X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}})-\hat{u}^2(T_{k+1},X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}})\\
&\qquad\qquad+\frac{1}{2}\Delta_{T_{k+1}}\left(f(\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}}))-f(\hat{\Theta}^{\pi,2}_{k+1})\right)\Big]\\
&\qquad+(1+\Delta_{T_k}\Psi_1)(\epsilon^s_{\hat{u}^2,k}(\pi)+\epsilon^c_{\hat{u}^2,k})(\pi)\\
&\qquad+\frac{1}{2}[\Psi_2\Delta_{T_k}+\Psi_2'\Delta_{T_{k+1}}][v(T_k,\hat{X}^\pi_{T_k})-\hat{v}^2(T_k,\hat{X}^\pi_{T_k})]\\
&\qquad+\Delta_{T_{k+1}}\frac{\Psi_1'}{2}[u(T_k,\hat{X}^\pi_{T_k})-\tilde{u}(T_k,\hat{X}^\pi_{T_k})]\\
&\qquad+\frac{1}{2}[\Psi_3\Delta_{T_k}+\Psi_3'\Delta_{T_{k+1}}](\langle\mu_{T_k},\varphi_f[\cdot,u(T_k,\cdot)]\rangle-\langle\hat{\mu}_{T_k},\varphi_f[\cdot,u(T_k,\cdot),\cdot]\rangle),\\
&\qquad+\Delta_{T_k}\frac{\Psi_3}{2}(\langle\hat{\mu}_{T_k},\varphi_f[\cdot,u(T_k,\cdot)]\rangle-\langle\hat{\mu}_{T_k},\varphi_f[\cdot,\hat{u}^2(T_k,\cdot),\cdot]\rangle)\\
&\qquad+\Delta_{T_{k+1}}\frac{\Psi_3'}{2}(\langle\hat{\mu}_{T_k},\varphi_f[\cdot,u(T_k,\cdot)]\rangle-\langle\hat{\mu}_{T_k},\varphi_f[\cdot,\tilde{u}(T_k,\cdot),\cdot]\rangle)
\end{aligned}
$$

Then, replacing the expansion for the error of $\hat{v}^2$ in terms of (2.6.31),(2.6.32), (2.6.30) and the one for $(u-\tilde{u})$ in terms of (2.6.33), (2.6.34), (2.6.35), and up to rescaling some of the $\Psi$ random variables, one obtains

$$
\begin{aligned}
&(u-\hat{u}^2)(T_k,\hat{X}^\pi_{T_k})+\frac{1}{2}\Delta_{T_k}\left(f(\Theta^\mu_k(\hat{X}^\pi_{T_k}))-f(\hat{\Theta}^{\pi,2}_k)\right) \qquad (2.6.42)\\
&=\mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\Big((1+\Delta_{T_{k+1}}\Psi_1''+\Delta_{T_k}\zeta_{k+1}\Psi_2'')\Big[(u-\hat{u}^2)(T_{k+1},X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}})\\
&\qquad+\frac{1}{2}\Delta_{T_{k+1}}\left(f(\Theta^\mu_{k+1}(X^{T_k,\hat{X}^\pi_{T_k},\mu}_{T_{k+1}}))-f(\hat{\Theta}^{\pi,2}_{k+1})\right)\Big]\Big)\\
&-\mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\left(\left[(u-\hat{u}^2)(T_{k+1},X^{T_k,\hat{X}^\pi_{T_k},\hat{\mu}}_{T_{k+1}})\right](\Delta_{T_{k+1}}\Psi_1''+\Delta_{T_{k+1}}\zeta_{k+1}\Psi_2'')\right)\\
&+(1+\Delta_{T_k}\Psi_1)(\epsilon^s_{\hat{u}^2,k}(\pi)+\epsilon^c_{\hat{u}^2,k}(\pi))+\Delta_{T_{k+1}}\Psi_3''(\epsilon^s_{\tilde{u},k}(\pi)+\epsilon^c_{\tilde{u},k}(\pi))\\
&+\Delta_{T_{k+1}}\Psi_4''(\epsilon^s_{\hat{v}^2,k}(\pi)+\epsilon^c_{\hat{v}^2,k}(\pi))\\
&+\frac{1}{2}[\Psi_3\Delta_{T_k}+\Psi_3'\Delta_{T_{k+1}}](\langle\mu_{T_k},\varphi_f[\cdot,u(T_k,\cdot)]\rangle-\langle\hat{\mu}_{T_k},\varphi_f[\cdot,u(T_k,\cdot),\cdot]\rangle),\\
&+\Delta_{T_k}\frac{\Psi_3}{2}(\langle\hat{\mu}_{T_k},\varphi_f[\cdot,u(T_k,\cdot)]\rangle-\langle\hat{\mu}_{T_k},\varphi_f[\cdot,\hat{u}^2(T_k,\cdot),\cdot]\rangle)\\
&+\Delta_{T_{k+1}}\frac{\Psi_3'}{2}(\langle\hat{\mu}_{T_k},\varphi_f[\cdot,u(T_k,\cdot)]\rangle-\langle\hat{\mu}_{T_k},\varphi_f[\cdot,\tilde{u}(T_k,\cdot),\cdot]\rangle).
\end{aligned}
$$

We are now in position to give the dynamics of the sum of the maximal errors $\mathcal{E}_u^2(k) + \Delta_{T_k}\mathcal{E}_f(k)$. We use Cauchy-Schwartz inequality on the first two terms of (2.6.42), bound the last two terms in (2.6.42) using the Lipschitz property of $\varphi_f$, as in the first order error analysis, we can deduce that for some constants $C, C'$,

$$\begin{aligned}\mathcal{E}_u^2(k) + \Delta_{T_{k+1}}\mathcal{E}_f(k) \quad &\leq \quad (1 + C\Delta_{T_{k+1}})\left[\mathcal{E}_u^2(k+1) + \Delta_{T_k}\mathcal{E}_f(k+1)\right] \\ &\quad + C'(\bar{\epsilon}_2(k) + \Delta_{T_{k+1}}N^{-[(m-1)\wedge q]/2}) \end{aligned} \tag{2.6.43}$$

with

$$\begin{aligned}\bar{\epsilon}_2(k) \quad &= \quad \sup_{\pi\in\mathcal{S}_k}|\epsilon^s_{\hat{u}^2,k}(\pi) + \Delta_{T_{k+1}}\epsilon^s_{\hat{v}^2,k}(\pi)| + \sup_{\pi\in\mathcal{S}_k}|\epsilon^c_{\hat{u}^2,k}(\pi) + \Delta_{T_{k+1}}\epsilon^c_{\hat{v}^2,k}(\pi)| \\ &\quad + \Delta_{T_{k+1}}\big(\sup_{\pi\in\mathcal{S}_k}|\epsilon^s_{\tilde{u},k}(\pi) + \epsilon^c_{\tilde{u},k}(\pi)|\big). \end{aligned} \tag{2.6.44}$$

We can bound the two first terms of $\bar{\epsilon}_2(k)$ using respectively Claims 2.6.7, 2.6.8, while the last term may be bounded using Claims 2.6.4 and 2.6.5 as we have discussed before, so that there is a constant $C''$

$$\begin{aligned}\bar{\epsilon}_2(k) \leq C''\Bigg(&\sup_{s\in[T_k,T_{k+1}]}||u(s,\cdot)||_{4,\infty}\Delta^3_{T_{k+1}} \\ &+ [||u(T_{k+1},\cdot)||_{3,\infty} + M_u(4,T_{k+1})]\left[\Delta^{q+1}_{T_{k+1}} + \Delta^2_{T_{k+1}}N^{-[(m-1)\wedge 2q]/2}\right] \\ &+ \sum_{i=m-2}^{m+1}[||u(T_{k+1},\cdot)||_{i,\infty} + M_u(i,T_{k+1}))]\Delta^{(i+1)/2}_{T_{k+1}} + \sum_{i=m+2}^{m+3}M_u(i,T_{k+1})\Delta^{(i+1)/2}_{T_{k+1}}\Bigg). \end{aligned} \tag{2.6.45}$$

Given that the initialization step is the first order scheme, Claims 2.6.4, 2.6.5 and 2.6.6 imply $\mathcal{E}_u^2(N-1) \leq KN^{-2}$ and $\Delta_{T_{N-1}}\mathcal{E}_f^2(N-1) \leq KN^{-2}$. Moreover, under **(SB)** , we have from Lemma 2.7.2 that for all $n \in \mathbb{N}^*$ there exists a constant $K$, depending on the regularity of $V_{0:d}$ and $\phi$, such that for all $k \in \{0, \cdots, N-1\}$,

$$M_u(n, T_{k+1}) + \sup_{s\in[T_k,T_{k+1}]}||u(s,\cdot)||_{n,\infty} \leq K.$$

An application of the discrete Gronwall lemma on the sum $\mathcal{E}_u^2(k) + \Delta_{T_{k+1}}\mathcal{E}_f(k)$, gives

$$\sup_{k\leq N-1}\mathcal{E}_u^2(k) + \Delta_{T_k}\mathcal{E}_f(k) \leq C\left(\mathcal{E}_u^2(N-1) + \Delta_{T_N}\mathcal{E}_f(N-1) + N^{-[(m-1)\wedge q]/2} + \sum_{i=0}^{N-2}\bar{\epsilon}_2(k)\right).$$

Using (2.6.45) we deduce that, if $m \geq 7$,where the bound is a consequence of the second assertion of Claim 2.6.8, we have

$$\sup_{k\leq N-1}\mathcal{E}_u^2(k) + \Delta_{T_k}\mathcal{E}_f(k) \leq CN^{-2}.$$

As for the first order case, the previous result together with the expansion of $v - \hat{v}^2$ given (2.6.31), (2.6.32), (2.6.30) and Cauchy-Schwartz inequality, implies

$$\Delta^{1/2}_{T_k}\mathcal{E}_v^2(k) \leq CN^{-2}.$$

This concludes the proof of assertions (2.2.3) and (2.2.4) in Theorem 2.2.1. □

## 2.7 Mathematical tools

Here we will intensively use the notions defined in section 2.5.

### *2.7.1* The conditional linear PDE

**Lemma 2.7.1.** *Let $\eta$ be a given family of probability measures on $\mathbb{R}^d$ and consider the PDE*

$$\left\{\begin{array}{l}\partial_t\psi(t,y)+\mathcal{L}^\eta\psi(t,y)=0, \text{ on } [0,T]\times\mathbb{R}^d\\ \psi(T,y)=\phi(y)\end{array}\right. \tag{2.7.1}$$

*where $\mathcal{L}^\eta$ is defined by (2.5.1). Suppose that assumption **(SB)** holds, then, this PDE admits a unique infinitely differentiable solution $\psi$ and for every multi-index $\beta\in\mathcal{M}$ there exists a positive constant $C$ depending on the regularity of $V_{0:d}$, $\varphi_{0:d}$ and $T$ such that, :*

$$||D_\beta\psi(t,.)||_\infty\le C||\phi||_{||\beta||,\infty} \tag{2.7.2}$$

*Proof.* Since the law that appears in the coefficients of the SDE is fixed, this is an obvious consequence of the regularity of the coefficients and the terminal condition. □

### *2.7.2* The conditional semi-linear PDE

**Lemma 2.7.2.** *Under **(SB)** there exists a function $u$ from $[0,T]\times\mathbb{R}^d$ to $\mathbb{R}$ such that*

$$Y_t^y=u(t,X_t^y)$$

*where $Y_t^y$ is defined in (2.0.7). This function is in $C^{1,2}([0,T]\times\mathbb{R}^d,\mathbb{R})$ and is the unique solution of the semi-linear PDE :*

$$\left\{\begin{array}{l}\partial_t u(t,y)+\mathcal{L}^\mu u(t,y)=f\left(t,y,u(t,y),(\mathcal{V}^\mu u(t,y))^T,\langle\mu_t,\varphi_f(\cdot,u(t,\cdot))\rangle\right), \text{ on } [0,T]\times\mathbb{R}^d\\ u(T,y)=\phi(y)\end{array}\right. \tag{2.7.3}$$

*where $\mathcal{L}^\mu$ is defined as in (2.0.10).*

*Moreover, $u$ is infinitely differentiable, and for every multi-index $\beta\in\mathcal{M}$ there exists a positive constant $C$ depending on the regularity of $V_{0:d}, f, \varphi_{0:d}$, $\phi$ and $T$ such that,*

$$||D_\beta u||_\infty\le C \tag{2.7.4}$$

**Proof of Lemma 2.7.2.**

(i) *Existence and PDE solution.* Consider the conditional BSDE :

$$\left\{\begin{array}{l}dX_s^{t,y,x}=\sum_{i=0}^d V_i(s,X_s^{t,y,x},\mathbb{E}\left[\varphi_i(X_s^x)\right])dB_s^i\\ d\bar{Y}_s^{t,y,x}=-f(s,X_s^{t,y,x},\bar{Y}_s^{t,y,x},\bar{Z}_s^{t,y,x},\mathbb{E}\left[\varphi_f(X_s^x,Y_s^x)\right])ds+\bar{Z}_s^{t,y,x}dB_s^{1:d}\\ X_t^{t,y,x}=y,\quad \bar{Y}_T^{t,y,x}=\phi(X_T^{t,y,x}),\end{array}\right. \tag{2.7.5}$$

for $s$ in $[t,T]$. Note that the McKean term in (2.7.5), $\mathbb{E}\varphi_f(X_s^x,Y_s^x)$, does not depend on $y$. In fact, if the solution of (2.0.7) is found, one can consider the term $\mathbb{E}\varphi_f(X_s^x,Y_s^x)$ simply as a term depending on time, so that conditionally to knowing the joint law of $(X_t^x,Y_t^x,\ 0\le t\le T)$, equation (2.7.5) is classical and Markov. As pointed out before, the existence of a unique solution to (2.0.7) follows the lines of the results in [BLP09].

It is clear from the previous discussion that we might apply classical results on BSDE to analyze equation (2.7.5). In particular, we have from the results of Pardoux and Peng in [PP92] given the regularity in space of $f$ and $\phi$, that the mapping $(t,y)\mapsto\bar{u}(t,y)=\bar{Y}_t^{t,y,x}$ the solution of (2.7.5) is differentiable, once in time and twice in space with bounded derivatives and satisfies the PDE :

$$\left\{\begin{array}{l}\partial_t\bar{u}(t,y)+\mathcal{L}^\mu\bar{u}(t,y)=f\left(t,y,\bar{u}(t,y),(\mathcal{V}^\mu\bar{u}(t,y))^T,\mathbb{E}[\varphi_f(X_t^x,Y_t^x)]\right)\\ \bar{u}(T,y)=\phi(y)\end{array}\right. ,$$

Besides, thanks to the bounds on $f$ and $\phi$ it follows from Proposition 3.2 in [BDH+03] that

$$||\bar{u}(t,\cdot)||_\infty \le C(T, V_{0:d}, f)(1 + ||\phi||_\infty^p)^{1/p} \le C(T, V_{0:d}, f, \phi).$$

Moreover, as shown in [PP92] the derivative of $\nabla_y \bar{Y}_s^{t,y,x}$ is given as the solution of the BSDE :

$$\begin{aligned}\nabla_y \bar{Y}_s^{t,y,x} =& \nabla_y \phi(X_T^{t,y,x}) \nabla_y X_T^{t,y,x} - \int_s^T \nabla_y \bar{Z}_r^{t,y,x} dB_r \\ &+ \int_s^T \Big[\partial_y f(r, X_r^{t,y,x}, \bar{Y}_r^{t,y,x}, \bar{Z}_r^{t,y,x}, \mathbb{E}[\varphi(X_r^x, Y_r^x, )]) \nabla_y X_r^{t,y,x} \\ &\qquad + \partial_{y'} f(r, X_r^{t,y,x}, \bar{Y}_r^{t,y,x}, \bar{Z}_r^{t,y,x}, \mathbb{E}[\varphi(X_r^x, Y_r^x)]) \nabla_y \bar{Y}_r^{t,y,x} \\ &\qquad + \partial_z f(r, X_r^{t,y,x}, \bar{Y}_r^{t,y,x}, \bar{Z}_r^{t,y,x}, \mathbb{E}[\varphi(X_r^x, Y_r^x)]) \nabla_y \bar{Z}_r^{t,y,x}\Big] dr,\end{aligned} \tag{2.7.6}$$

and $\nabla_y \bar{u}(t,y) = \nabla_y \bar{Y}_t^{t,y,x}$. Once again using the bounds on the derivatives in space of $\phi$ and $f$ and the results in [BDH+03] we deduce

$$||\nabla_y \bar{u}(t,\cdot)||_\infty \le C(T, V_{0:d}, f)(1 + ||\partial_y \phi||_\infty^p)^{1/p} \le C(T, V_{0:d}, f, \phi). \tag{2.7.7}$$

Finally, one can show that $\bar{u}$ solves (2.7.3). To see this, notice that due to the uniqueness of the solutions to (2.0.7) and (2.7.5),

$$\bar{u}(t, X_t^x) = \bar{u}(t, X_t^{0,x,x}) = \bar{Y}_t^{t, X_t^{0,x}, x} = \bar{Y}_t^{0,x,x} = Y_t^x.$$

This equality implies

$$\mathbb{E}[\varphi_f(X_s^x, \bar{u}(s, X_s^x))] = \mathbb{E}[\varphi_f(X_s^x, Y_s^x)]$$

for all $s$ in $[0,T]$. Therefore we set $u := \bar{u}$. This concludes the proof of the first assertion.

(ii) *Control on the derivatives.*
To prove the regularity of $u$ and the the bound on its derivatives, we consider first the case involving only space derivatives. In this case the whole argument of Pardoux and Peng may be iterated reasoning on the BSDE for the first derivative, as long as the hypotheses remain valid, to obtain a BSDE for higher order derivatives in space. We turn the reader to the paper of Crisan and Delarue [CD12c] where this is done in detail (taking into account the additional law dependence that must be considered in our framework).

It remains to consider the case of general derivatives including time derivatives. As we have said before, iterative applications of the Pardoux and Peng argument lead to PDEs similar to (2.7.3). Then, we can argue that we are able to differentiate once in time for every two derivatives in space. It is also clear that the control on the space derivatives plus the regularity properties of the coefficients imply the control for time derivatives.

### *2.7.3* One-step errors

Let $\eta$ be a given family of probability measures and $X^{t,y,\eta}$, $\tilde{X}^{t,y,\eta}$ defined as in (2.5.2) and (2.5.3). We recall that $\hat{\mu}$ denotes the discrete probability measure defined by Algorithm 1 and $\mu$ the law of the forward part of (2.0.7).

**Lemma 2.7.3.** *Let $g$ be a $C_b^2$ function from $\mathbb{R}^d$ to $\mathbb{R}$. Then, there exists a constant $C$ depending only on $V_{0:d}$ and $T$ such that for all $k = 1, \cdots, N-1$ :*

$$\begin{aligned}
&\left|(P^{\eta}_{T_k,T_{T_{k+1}}} - \tilde{P}^{\hat{\mu}}_{T_k,T_{T_{k+1}}})g(y)\right| \quad (2.7.8)\\
&\leq C||g||_{2,\infty} \sum_{i=0}^{d} \int_{T_k}^{T_{k+1}} \left| \langle \eta_t, \varphi_i \rangle - \sum_{p=0}^{q-1} [(t-T_k)^p/p!] \langle \eta_{T_k}, (\mathcal{L}^{\eta})^p \varphi_i \rangle \right| dt \\
&+ C||g||_{2,\infty} \sum_{i=0}^{d} \sum_{p=0}^{q-1} \Delta_{T_{k+1}}^{p+1} \left| \langle \eta_{T_k}, (\mathcal{L}^{\eta})^p \varphi_i \rangle - \langle \hat{\mu}_{T_k}, (\mathcal{L}^{\hat{\mu}})^p \varphi_i \rangle \right|.
\end{aligned}$$

*Moreover, if $g$ is $C_b^3$, there exists a constant $C$ depending only on $V_{0:d}, d$ such that for all $l = 1, \cdots, d$ and $k = 1, \cdots, N-1$ :*

$$\begin{aligned}
&\left|\mathbb{E}\left(\left[g\left(X^{T_k,y,\eta}_{T_{k+1}}\right) - g\left(\tilde{X}^{T_k,y,\hat{\mu}}_{T_{k+1}}\right)\right] I^{T_k,T_{k+1}}_{(l)}\right)\right| \quad (2.7.9)\\
&\leq C||g||_{3,\infty} \sum_{i=0}^{d} \int_{T_k}^{T_{k+1}} \int_{T_k}^{t} \left| \langle \eta_s, \varphi_i \rangle - \sum_{p=0}^{q-1} [(t-T_k)^p/p!] \langle \eta_{T_k}, (\mathcal{L}^{\eta})^p \varphi_i \rangle \right| dsdt \\
&+ C||g||_{3,\infty} \sum_{i=0}^{d} \sum_{p=0}^{q-1} \Delta_{T_{k+1}}^{p+1} \left| \langle \eta_{T_k}, (\mathcal{L}^{\eta})^p \varphi_i \rangle - \langle \hat{\mu}_{T_k}, (\mathcal{L}^{\hat{\mu}})^p \varphi_i \rangle \right|
\end{aligned}$$

*and*

$$\begin{aligned}
&\left|\mathbb{E}\left(\left[g\left(X^{T_k,y,\eta}_{T_{k+1}}\right) - g\left(\tilde{X}^{T_k,y,\hat{\mu}}_{T_{k+1}}\right)\right] I^{T_k,T_{k+1}}_{(0,l)}\right)\right| \quad (2.7.10)\\
&\leq C||g||_{3,\infty} \Delta_{T_{k+1}} \sum_{i=0}^{d} \int_{T_k}^{T_{k+1}} \int_{T_k}^{t} \left| \langle \eta_s, \varphi_i \rangle - \sum_{p=0}^{q-1} [(t-T_k)^p/p!] \langle \eta_{T_k}, (\mathcal{L}^{\eta})^p \varphi_i \rangle \right| dsdt \\
&+ C||g||_{3,\infty} \sum_{i=0}^{d} \sum_{p=0}^{q-1} \Delta_{T_{k+1}}^{p+2} \left| \langle \eta_{T_k}, (\mathcal{L}^{\eta})^p \varphi_i \rangle - \langle \hat{\mu}_{T_k}, (\mathcal{L}^{\hat{\mu}})^p \varphi_i \rangle \right|.
\end{aligned}$$

**Lemma 2.7.4.** *Let $n \geq 1$, $g$ be a $C_b^{n+2}$ function from $\mathbb{R}^d$ to $\mathbb{R}$ and $\mathbb{Q}$ be a cubature measure of order $n$. Then, there exist constants $C, C'$ depending only on $V_{0:d}, d, n$ such that for all $i = 1, \ldots, d$, for all $k = 1, \cdots, N-1$ :*

$$\begin{aligned}
\left|(\tilde{P}^{\eta}_{T_k,T_{k+1}} - \tilde{Q}^{\eta}_{T_k,T_{k+1}})g(y)\right| &= \left|(\mathbb{E} - \mathbb{E}_{\mathbb{Q}})\left[g(\tilde{X}^{T_k,y,\eta}_{T_{k+1}})\right]\right| \\
&\leq C \sum_{l=n+1}^{n+2} ||g||_{l,\infty} \Delta_{T_{k+1}}^{(l)/2} \quad (2.7.11)
\end{aligned}$$

$$\left|(\mathbb{E} - \mathbb{E}_{\mathbb{Q}})\left[g(\tilde{X}^{T_k,y,\eta}_{T_{k+1}}) I^{T_k,T_{k+1}}_{(i)}\right]\right| \leq C \sum_{l=n}^{n+1} ||g||_{l,\infty} \Delta_{T_{k+1}}^{(l+1)/2} \quad (2.7.12)$$

$$\left|(\mathbb{E} - \mathbb{E}_{\mathbb{Q}})\left[g(\tilde{X}^{T_k,y,\eta}_{T_{k+1}}) I^{T_k,T_{k+1}}_{(0,i)}\right]\right| \leq C' \sum_{l=n-2}^{n-1} ||g||_{l,\infty} \Delta_{T_{k+1}}^{(l+3)/2}, \quad (2.7.13)$$

*for all $y \in \mathbb{R}^d$.*

**Lemma 2.7.5.** *Let $n \geq 1$ and $g$ be a $C_b^{n+2}$ function from $\mathbb{R}^d$ to $\mathbb{R}$. Let $\mathbb{Q}$ be a cubature measure of order $n$, and $\hat{X}$ be the associated cubature tree. Then, there exists a constant $C$ depending only on $d$, $q$, $V_{0:d}, n, T, ||\varphi_{0:d}||_{2q+n+2,\infty}$ such that, for all $k = 1, \cdots, N-1$ :*

$$\begin{aligned} & \left|\mathbb{E}\left[g\left(X_{T_{k+1}}^{T_k,y,\mu}\right)\right] - \mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\left[g\left(X_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right]\right| \qquad (2.7.14) \\ & \leq C||g||_{2,\infty}\left[\Delta_{T_{k+1}}^{q+1} + \Delta_{T_{k+1}} N^{-[(n-1)\wedge 2q]/2}\right] + C||g||_{n+1,\infty}\Delta_{T_{k+1}}^{(n+1)/2} + C||g||_{n+2,\infty}\Delta_{T_{k+1}}^{(n+2)/2}, \end{aligned}$$

$$\begin{aligned} & \left|\mathbb{E}\left[\left(g\left(X_{T_{k+1}}^{T_k,y,\mu}\right) - g\left(\hat{X}_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right) I_{(l)}^{T_k,T_{k+1}}\right]\right| \qquad (2.7.15) \\ & \leq C||g||_{3,\infty}\left[\Delta_{T_{k+1}}^{q+1} + \Delta_{T_{k+1}}^{2} N^{-[(n-1)\wedge 2q]/2}\right] + C||g||_{n,\infty}\Delta_{T_{k+1}}^{(n+1)/2} + C||g||_{n+1,\infty}\Delta_{T_{k+1}}^{(n+2)/2} \end{aligned}$$

*and*

$$\begin{aligned} & \left|\mathbb{E}\left[\left(g\left(X_{T_{k+1}}^{T_k,y,\mu}\right) - g\left(\hat{X}_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right) I_{(0,l)}^{T_k,T_{k+1}}\right]\right| \qquad (2.7.16) \\ & \leq C||g||_{3,\infty}\left[\Delta_{T_{k+1}}^{q+2} + \Delta_{T_{k+1}}^{3} N^{-[(n-1)\wedge 2q]/2}\right] + C||g||_{n-2,\infty}\Delta_{T_{k+1}}^{(n+1)/2} + C||g||_{n-1,\infty}\Delta_{T_{k+1}}^{(n+2)/2}. \end{aligned}$$

### *2.7.4* Proofs of Lemmas 2.7.3, 2.7.4 and 2.7.5

***2.7.4.1 Proof of Lemma 2.7.3*** Let $k \in \{0, \cdots, N-1\}$, consider the PDE (2.7.1) with $g$ as boundary condition. It is clear that this PDE admits a unique solution $\tilde{u}$ and that there exists a positive constant $C(T, V_{0:d})$ such that for every multi-index of space derivatives with $\beta \in \mathcal{A}_3$

$$||\tilde{u}||_\infty + ||D_\beta \tilde{u}||_\infty \leq C(T, V_{0:d})||g||_{||\beta||,\infty}. \qquad (2.7.17)$$

Let us write :

$$\left|(P^{\eta}_{T_k,T_{T_{k+1}}} - \tilde{P}^{\hat{\mu}}_{T_k,T_{T_{k+1}}})g(y)\right| = \left|\mathbb{E}\left[g\left(X_{T_{k+1}}^{T_k,y,\eta}\right) - g\left(\tilde{X}_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right]\right|.$$

Now, we have

$$\begin{aligned} g\left(X_{T_{k+1}}^{T_k,y,\eta}\right) &= \tilde{u}(T_k, y) + \int_{T_k}^{T_{k+1}} \mathcal{V}^{\eta}_{(0)}\tilde{u}(s, X_s^{T_k,y,\eta})ds + \sum_{j=1}^{d}\int_{T_k}^{T_{k+1}} \mathcal{V}^{\eta}_{(j)}\tilde{u}(s, X_s^{T_k,y,\eta})dB_s^j \\ &= \tilde{u}(T_k, y) + \sum_{j=1}^{d}\int_{T_k}^{T_{k+1}} \mathcal{V}^{\eta}_{(j)}\tilde{u}(s, X_s^{T_k,y,\eta})dB_s^j, \qquad (2.7.18) \end{aligned}$$

using the fact that $\tilde{u}$ is the solution of (2.7.1), and

$$\begin{aligned} g\left(\tilde{X}_{T_{k+1}}^{T_k,y,\hat{\mu}}\right) &= \tilde{u}(T_k, y) + \int_{T_k}^{T_{k+1}} \mathcal{V}^{\hat{\mu}}_{(0)}\tilde{u}(s, \tilde{X}_s^{T_k,y,\hat{\mu}})ds + \sum_{j=0}^{d}\int_{T_k}^{T_{k+1}} \mathcal{V}^{\hat{\mu}}_{(j)}\tilde{u}(s, X_s^{T_k,y,\hat{\mu}})dB_s^j \\ &= \tilde{u}(T_k, y) + \int_{T_k}^{T_{k+1}} (\mathcal{L}^{\eta} - \tilde{\mathcal{L}}^{\hat{\mu}})\tilde{u}(s, \tilde{X}_s^{T_k,y,\hat{\mu}})ds \\ &\quad + \sum_{j=1}^{d}\int_{T_k}^{T_{k+1}} \mathcal{V}^{\hat{\mu}}_{(j)}\tilde{u}(s, \tilde{X}_s^{T_k,y,\hat{\mu}})dB_s^j. \qquad (2.7.19) \end{aligned}$$

Therefore,

$$\mathbb{E}\left[g\left(X_{T_{k+1}}^{T_k,y,\eta}\right) - g\left(\tilde{X}_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right] = \mathbb{E}\int_{T_k}^{T_{k+1}} \left(\mathcal{L}^{\eta} - \tilde{\mathcal{L}}^{\hat{\mu}}\right)\tilde{u}(t, \tilde{X}_t^{T_k,y,\hat{\mu}})dt, \qquad (2.7.20)$$

since $\partial_t \tilde{u} = -\mathcal{L}^\eta \tilde{u}$. As (2.7.17) implies that $\tilde{u}$ and its two first derivatives are bounded, we may control the term above by the difference between the two generators, then by the difference between the frozen (in space of probability measure) $(\eta_t)_{T_K \le t \le T_{k+1}}$ and the approximate and frozen (in time) measure $\hat{\mu}_{T_k}$. Hence, taking into account the particular dependence on the measure in our framework, we deduce :

$$
\begin{aligned}
&\left| \mathbb{E}\left[ g\left( X^{T_k,y,\eta}_{T_{k+1}} \right) - g\left( \tilde{X}^{T_k,y,\hat{\mu}}_{T_{k+1}} \right) \right] \right| \\
&\le C||g||_{2,\infty} \int_{T_k}^{T_{k+1}} \left[ \sum_{i=0}^{d} \left| \langle \eta_t, \varphi_i \rangle - \sum_{p=0}^{q-1} [(t-T_k)^p/p!] \langle \eta_{T_k}, (\mathcal{L}^\eta)^p \varphi_i \rangle \right| \right. \\
&\qquad \left. + \sum_{i=0}^{d} \sum_{p=0}^{q-1} \Delta^p_{T_{k+1}} \left| \langle \eta_{T_k}, (\mathcal{L}^\eta)^p \varphi_i \rangle - \langle \hat{\mu}_{T_k}, (\mathcal{L}^{\hat{\mu}})^p \varphi_i \rangle \right| \right] dt \\
&\le C(T, V_{0:d}) ||g||_{2,\infty} \int_{T_k}^{T_{k+1}} \sum_{i=0}^{d} \left| \langle \eta_t, \varphi_i \rangle - \sum_{p=0}^{q-1} [(t-T_k)^p/p!] \langle \eta_{T_k}, (\mathcal{L}^\eta)^p \varphi_i \rangle \right| dt \\
&\quad + C(T, V_{0:d}) ||g||_{2,\infty} \sum_{i=0}^{d} \sum_{p=0}^{q-1} \Delta^p_{T_{k+1}} \Delta_{T_{k+1}} \left| \langle \eta_{T_k}, (\mathcal{L}^\eta)^p \varphi_i \rangle - \langle \hat{\mu}_{T_k}, (\mathcal{L}^{\hat{\mu}})^p \varphi_i \rangle \right|
\end{aligned}
$$

This concludes the proof of the first assertion. Now, we deduce from (2.7.18) and (2.7.19) and integration by parts, that :

$$
\begin{aligned}
&\mathbb{E}\left( \left[ g\left( X^{T_k,y,\eta}_{T_{k+1}} \right) - g\left( \tilde{X}^{T_k,y,\hat{\mu}}_{T_{k+1}} \right) \right] I^{T_k,T_{k+1}}_{(l)} \right) \qquad (2.7.21) \\
&= \mathbb{E} \int_{T_k}^{T_{k+1}} \left( \mathcal{L}^\eta - \tilde{\mathcal{L}}^{\hat{\mu}} \right) \tilde{u}(t, \tilde{X}^{T_k,y,\hat{\mu}}_t) I^{T_k,t}_{(l)} dt \\
&\quad + \mathbb{E} \int_{T_k}^{T_{k+1}} \left[ \mathcal{V}_{(l)}(t, X^{T_k,y,\eta}_t, \langle \eta_t, \varphi_l \rangle) - \mathcal{V}_{(l)}(t, X^{T_k,y,\eta}_t, \langle \hat{\mu}_{T_k}, \varphi_l \rangle) \right] \tilde{u}\left( t, X^{T_k,y,\eta}_t \right) dt \\
&\quad + \frac{1}{2} \mathbb{E} \int_{T_k}^{T_{k+1}} \left\{ \left[ \mathcal{V}_{(l)}(t, X^{T_k,y,\eta}_t, \langle \hat{\mu}_{T_k}, \varphi_l \rangle) \tilde{u}(t, X^{T_k,y,\eta}_t) \right. \right. \qquad (2.7.22) \\
&\quad \left. \left. - \mathcal{V}_{(l)}(t, \tilde{X}^{T_k,y,\hat{\mu}}_t, \langle \hat{\mu}_{T_k}, \varphi_l \rangle) \tilde{u}(t, \tilde{X}^{T_k,y,\hat{\mu}}_t) \right] \right\} dt.
\end{aligned}
$$

On a first hand, note that the first two terms in the right hand side above may be controlled by the difference between the coefficients times the supremum norm of the first and second order derivatives of $\tilde{u}$ times the order of the integrals, as we did for (2.7.20). On the other hand, note that the first assertion of Lemma 2.7.3 can be applied to the function $\mathcal{V}_{(l)}(t, ., \langle \hat{\mu}_{T_k}, \varphi_l \rangle) \tilde{u}(t, .)$ in the last term on the right hand side above. These arguments, together with the bound (2.7.17) lead to :

$$
\begin{aligned}
&\left|\mathbb{E}\left(\left[g\left(X_{T_{k+1}}^{T_k,y,\eta}\right)-g\left(\tilde{X}_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right]I_{(l)}^{T_k,T_{k+1}}\right)\right| \\
&\leq C(T,V_{0:d})\int_{T_k}^{T_{k+1}}\Bigg\{(||g||_{1,\infty}+||g||_{2,\infty}(t-T_k)^{1/2}) \\
&\quad\times\sum_{i=0}^{d}\left|\langle\eta_t,\varphi_i\rangle-\sum_{p=0}^{q-1}[(t-T_k)^p/p!]\langle\eta_{T_k},(\mathcal{L}^\eta)^p\varphi_i\rangle\right|\Bigg\}dt \\
&\quad+C(T,V_{0:d})\sum_{i=0}^{d}\sum_{p=0}^{q-1}\Delta_{T_{k+1}}^p\Delta_{T_{k+1}}(||g||_{1,\infty}+||g||_{2,\infty}\Delta_{T_{k+1}}^{1/2})\left|\langle\eta_{T_k},(\mathcal{L}^\eta)^p\varphi_i\rangle-\langle\hat{\mu}_{T_k},(\mathcal{L}^{\hat{\mu}})^p\varphi_i\rangle\right| \\
&\quad+C(T,V_{0:d})||g||_{3,\infty}\Bigg\{\int_{T_k}^{T_{k+1}}\int_{T_k}^{t}\sum_{i=0}^{d}\left|\langle\eta_s,\varphi_i\rangle-\sum_{p=0}^{q-1}[(s-T_k)^p/p!]\langle\eta_{T_k},(\mathcal{L}^\eta)^p\varphi_i\rangle\right|dsdt \\
&\qquad\qquad+\sum_{i=0}^{d}\sum_{p=0}^{q-1}\Delta_{T_{k+1}}^p\Delta_{T_{k+1}}^2\left|\langle\eta_{T_k},(\mathcal{L}^\eta)^p\varphi_i\rangle-\langle\hat{\mu}_{T_k},(\mathcal{L}^{\hat{\mu}})^p\varphi_i\rangle\right|\Bigg\},
\end{aligned}
$$

and this concludes the proof of the second assertion. Finally, (2.7.18) and (2.7.19) and integration by parts, give :

$$
\begin{aligned}
&\mathbb{E}\left(\left[g\left(X_{T_{k+1}}^{T_k,y,\eta}\right)-g\left(\tilde{X}_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right]I_{(0,l)}^{T_k,T_{k+1}}\right) \\
&=\mathbb{E}\int_{T_k}^{T_{k+1}}\left(\mathcal{L}^\eta-\tilde{\mathcal{L}}^{\hat{\mu}}\right)\tilde{u}(t,\tilde{X}_t^{T_k,y,\eta})I_{(0,l)}^{T_k,t}dt \\
&\quad+\mathbb{E}\int_{T_k}^{T_{k+1}}\left[\mathcal{V}_{(l)}(t,X_t^{T_k,y,\eta},\langle\eta_t,\varphi_l\rangle)-\mathcal{V}_{(l)}(t,X_t^{T_k,y,\eta},\langle\hat{\mu}_{T_k},\varphi_l\rangle)\right]\tilde{u}\left(t,X_t^{T_k,y,\eta}\right)I_{(0)}^{T_k,t}dt \\
&\quad+\frac{1}{2}\mathbb{E}\int_{T_k}^{T_{k+1}}\Big\{\big[\mathcal{V}_{(l)}(t,X_t^{T_k,y,\eta},\langle\hat{\mu}_{T_k},\varphi_l\rangle)\tilde{u}(t,X_t^{T_k,y,\eta}) \\
&\quad-\mathcal{V}_{(l)}(t,\tilde{X}_t^{T_k,y,\hat{\mu}},\langle\hat{\mu}_{T_k},\varphi_l\rangle)\tilde{u}(t,\tilde{X}_t^{T_k,y,\hat{\mu}})\big]I_{(0)}^{T_k,t}\Big\}dt.
\end{aligned}
$$

Note the similarity with (2.7.21). So that a similar development gives

$$
\begin{aligned}
&\left|\mathbb{E}\left(\left[g\left(X_{T_{k+1}}^{T_k,y,\eta}\right)-g\left(\tilde{X}_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right]I_{(0,l)}^{T_k,T_{k+1}}\right)\right| \\
&\leq C(T,V_{0:d})\int_{T_k}^{T_{k+1}}\left(||g||_{1,\infty}(t-T_k)+||g||_{2,\infty}(t-T_k)^{3/2}\right) \\
&\quad\times\sum_{i=0}^{d}\left|\langle\eta_t,\varphi_i\rangle-\sum_{p=0}^{q-1}[(t-T_k)^p/p!]\langle\eta_{T_k},(\mathcal{L}^\eta)^p\varphi_i\rangle\right|dt \\
&\quad+C(T,V_{0:d})\Delta_{T_{k+1}}\sum_{i=0}^{d}\sum_{p=0}^{q-1}\Delta_{T_{k+1}}^p\Delta_{T_{k+1}}\left(||g||_{1,\infty}+||g||_{2,\infty}\Delta_{T_{k+1}}^{1/2}\right) \\
&\quad\times\left|\langle\eta_{T_k},(\mathcal{L}^\eta)^p\varphi_i\rangle-\langle\hat{\mu}_{T_k},(\mathcal{L}^{\hat{\mu}})^p\varphi_i\rangle\right| \\
&\quad+C(T,V_{0:d})\Delta_{T_{k+1}}||g||_{3,\infty}\left\{\int_{T_k}^{T_{k+1}}\int_{T_k}^{t}\sum_{i=0}^{d}\left|\langle\eta_s,\varphi_i\rangle-\sum_{p=0}^{q-1}[(s-T_k)^p/p!]\langle\eta_{T_k},(\mathcal{L}^\eta)^p\varphi_i\rangle\right|dsdt\right. \\
&\quad\left.+\sum_{i=0}^{d}\sum_{p=0}^{q-1}\Delta_{T_{k+1}}^p\Delta_{T_{k+1}}^2\left|\langle\eta_{T_k},(\mathcal{L}^\eta)^p\varphi_i\rangle-\langle\hat{\mu}_{T_k},(\mathcal{L}^{\hat{\mu}})^p\varphi_i\rangle\right|\right\}
\end{aligned}
$$

from where the last claim is deduced. □

***2.7.4.2 Proof of Lemma 2.7.4*** Let $k\in\{0,\cdots,N-1\}$, once again, we consider the unique infinitely differentiable solution $\tilde{u}$ of PDE (2.7.1) with $g$ as boundary condition. Recall that for every $\beta\in\mathcal{M}$ there exists a positive constant $C(T,V_{0:d})$ such that :

$$
||\tilde{u}||_\infty+||D_\beta\tilde{u}||_\infty\leq C(T,V_{0:d})||g||_{||\beta||,\infty}. \tag{2.7.23}
$$

The result then follows from Stratonovich-Taylor expansion of $(t,y)\mapsto\tilde{u}(t,y)$ around $(T_k,X_{T_k})$ by Theorem 5.6.1 in [KP92] and bounding the remainder as in Proposition 2.1 of [LV04]. □

***2.7.4.3 Proof of Lemma 2.7.5*** Note that

$$
\begin{aligned}
&\left|\mathbb{E}\left[g\left(X_{T_{k+1}}^{T_k,y,\mu}\right)\right]-\mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\left[g\left(X_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right]\right| \\
&=\left|\mathbb{E}\left[g\left(X_{T_{k+1}}^{T_k,y,\mu}\right)\right]-\mathbb{E}\left[g\left(\tilde{X}_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right]\right| \\
&\quad+\left|\mathbb{E}\left[g\left(\tilde{X}_{T_{k+1}}^{T_k,y,\mu}\right)\right]-\mathbb{E}_{\mathbb{Q}_{T_k,T_{k+1}}}\left[g\left(X_{T_{k+1}}^{T_k,y,\hat{\mu}}\right)\right]\right|.
\end{aligned} \tag{2.7.24}
$$

Combining estimate (2.7.8) of Lemma 2.7.3 with Claim 2.6.3 and (2.2.2) in Theorem 2.2.1, we get that the first term in the right hand side is bounded by :

$$
C||g||_{2,\infty}\left[\Delta_{T_{k+1}}^{q+1}+\Delta_{T_{k+1}}N^{-[(n-1)\wedge 2q]/2}\right].
$$

The second term in the right hand side of (2.7.24) can be estimated by combining this bound with the estimate (2.7.11) in 2.7.4 (when choosing $\eta=\mu$).

The other assertion follows from the same procedure, substituting (2.7.9) (resp. (2.7.10)) to (2.7.8) and (2.7.12) (resp. (2.7.13)) to (2.7.11).

## *2.8* Proofs of Corollary 2.2.2 and 2.2.3

### *2.8.1* Proof of Corollary 2.2.2

Many practical applications, particularly in finance, require the algorithm to be able to solve problems in which the boundary condition $\phi$ is less regular, e.g. when $\phi$ is just Lipschitz. In this section, we prove how the results obtained in the regular case extend to the case when assumption **(LB)** holds as Corollary 2.2.2 state.

**A preliminary result.** We use in addition an auxiliary result shown in the proof of Theorem 8 in [CG07] :

**Lemma 2.8.1.** *There exists a positive constant $C$ such that :*

$$\sum_{j=0}^{N-2} \Delta_{T_{j+1}}^{(m+1)/2} (T - T_j)^{-m/2} \leq CL(\gamma, m),$$

*where $L$ is defined in (2.2.6).*

We are now ready to examine the error convergence for the forward and backward components of the algorithm.

**Proof of the forward approximation in Corollary 2.2.2**

The regularity of the solution of the linear associated linear PDE is essential to our analysis. We start by stating a result in this sense under **(LB)**. This is summarized by

**Claim 2.8.2.** *Under **(LB)**, there exists a unique solution $\psi$ to the PDE (2.7.1) and for every multi-index $\beta \in \mathcal{M}$ there exists a constant $C$ such that :*

$$||D_\beta \psi(t, \cdot)||_\infty \leq C(T-t)^{-(||\beta||-1)/2} \tag{2.8.1}$$

*Proof.* This follows from classical results of parabolic equations with parameter, see Chapter 9, Section 3 of [Fri08].

□

Thanks to the uniform ellipticity assumption, even if the terminal condition is not differentiable, we know that the solution of the PDE (2.6.2) is smooth except at the boundary. Precisely, the gradient bounds (2.6.8) are now given by

$$||\nabla_y^n \psi(t, \cdot)||_\infty \leq C(T, V_{0:d}) ||\phi||_{1,\infty} (T-t)^{(1-n)/2}, \tag{2.8.2}$$

where $\psi$ is defined in (2.6.1). With this in hand, we can follow the proof exactly as the one of the corresponding forward part in Theorem 2.2.1 up to estimate (2.6.7) but where we separate the error on the last step, since there is no smoothing effect there. Then, plugging estimate (2.8.2) in (2.6.7) instead of (2.6.8), we get :

$$\begin{aligned}
&\left|(P_{T_0,T_N} - Q^{\hat{\mu}}_{T_0,T_N})\phi(x)\right| \\
&\leq C(T, V_{0:d})||\phi||_{1,\infty} \sum_{j=0}^{N-2}\sum_{p=0}^{q-1} \Delta^{p+1}_{T_{j+1}}(T-T_{j+1})^{-1/2}\sum_{i=0}^{d}\left|(P_{T_0,T_N} - Q^{\hat{\mu}}_{T_0,T_N})\varphi_i(x)\right| \\
&\quad + C(T, V_{0:d}, d)||\phi||_{1,\infty}\|\varphi\|_{2q,\infty}\sum_{j=0}^{N-2}\Delta^{q+1}_{T_{j+1}}(T-T_{j+1})^{-1/2} \\
&\quad + C(V_{0:d}, d, m)||\phi||_{1,\infty}\sum_{j=0}^{N-2}\sum_{l=m+1}^{m+2}\Delta^{\frac{l}{2}}_{T_{j+1}}(T-T_{j+1})^{(1-l)/2} \\
&\quad + \left|(P_{T_0,T_N} - P^{\eta}_{T_0,T_N})\phi(x)\right|.
\end{aligned}$$

We conclude the proof by using Lemma 2.8.1 on the sums and by combining Lipschitz property of $\phi$ and adapted time-step on the last step error. $\square$

### Proof of the backward approximation in Corollary 2.2.2

Just as in the forward case, our analysis relies on the regularization properties of the associated non-linear PDE under **(LB)**. We have

**Claim 2.8.3.** *Under **(LB)**, there exists a unique solution $u$ of (2.7.3), for all $(t,y) \in [0,T] \times \mathbb{R}^d$, it is given by*

$$u(t,y) = Y^{t,y,\mu}_t,$$

*where $Y^{t,y,\mu}_t$ is defined in (2.0.11). Moreover, for $\phi$ Lipschitz and bounded, for every multi-index $\beta \in \mathcal{M}$ there exists a positive constant $C$ depending on the regularity of $V_{0:d}, f, \varphi$ and $T$ such that :*

$$||D_\beta u(t,\cdot)||_\infty \leq C||\phi||^{|\beta|}_{1,\infty}(T-t)^{-(||\beta||-1)/2} \tag{2.8.3}$$

*Proof.* To prove Claim 2.8.3, we follow the same arguments given for Lemma 2.7.2. First, due to the regularity properties of the diffusion under the elliptic case, we have similar properties as those used in the paper of Crisan and Delarue [CD12c], even in the non-homogeneous case (notably, the integration by parts property as shown in [Ma02]). Hence, we get the control on derivatives result for space derivatives, and extend it, as before, to time derivatives. $\square$

Armed with the regularity of the function $u$, we can repeat the proof of the backward approximation in Theorem (2.2.1). We recover (2.6.23) for the first order scheme and (2.6.43) for the second order scheme, i.e.

$$\mathcal{E}^1_u(k) \leq \left(1 + C\Delta_{T_{k+1}}\right)\mathcal{E}^1_u(k+1) + \bar{\epsilon}(k+1), \tag{2.8.4}$$

$$\mathcal{E}^2_u(k) + \Delta_{T_{k+1}}\mathcal{E}_f(k) \leq (1 + C\Delta_{T_{k+1}})\left[\mathcal{E}^2_u(k+1) + \Delta_{T_k}\mathcal{E}_f(k+1)\right] + C'(\bar{\epsilon}_2(k) + \Delta_{T_{k+1}}N^{-[(m-1)\wedge q]/2}), \tag{2.8.5}$$

where $\bar{\epsilon}, \bar{\epsilon}_2$ are respectively defined in (2.6.24), (2.6.44). Now, if we show that

$$\sum_{k=0}^{N-1}\bar{\epsilon}(k) \leq N^{-1}; \quad \text{and} \quad \sum_{k=0}^{N-2}\bar{\epsilon}_2(k) \leq N^{-2}; \tag{2.8.6}$$

then, as in the smooth setting, we can apply Gronwall lemma on (2.8.4) and (2.8.5) and conclude on the desired rates of convergence for the approximation of $u$. The arguments for the rate of the approximation of $v$ are exactly as in the smooth setting thus completing the proof of the claimed result.

Therefore, we only need to prove (2.8.6). But, Claim 2.8.3 and the definition of $M_u$ given in (2.6.10) imply

$$M_u(n, T_{k+1}) \leq (T - T_{k+1})^{(1-n)/2} ||\phi||_{1,\infty}^n,$$

which together with Claim 2.8.3 show that under the Lipschitz boundary setup,

$$\begin{aligned}
\bar{\epsilon}(k+1) \quad &\leq \quad C\Bigg( (T-T_k)^{-1/2} \Delta_{T_{k+1}} \left[ \Delta^3_{T_{k+1}} + \Delta^q_{T_{k+1}} + N^{-[(m-1)\wedge 2q]/2} \right] \\
&\qquad + (T-T_k)^{-m/2} \Delta^{(m+1)/2}_{T_{k+1}} + (T-T_k)^{-(m+1)/2} \Delta^{(m+2)/2}_{T_{k+1}} \\
&\qquad + (T-T_k)^{-1} \Delta^{3/2}_{T_{k+1}} \left[ \Delta^{1/2}_{T_{k+1}} + \Delta^{q-1/2}_{T_{k+1}} + \Delta^{1/2}_{T_{k+1}} N^{-[(m-1)\wedge 2q]/2} \right] \Bigg). \\
&\leq \quad C\Bigg( (T-T_k)^{-1/2} \Delta_{T_{k+1}} N^{-1} + (T-T_k)^{-m/2} \Delta^{(m+1)/2}_{T_{k+1}} \\
&\qquad + (T-T_k)^{-(m+1)/2} \Delta^{(m+2)/2}_{T_{k+1}} + (T-T_k)^{-1} \Delta^{3/2}_{T_{k+1}} N^{-1/2} \Bigg);
\end{aligned} \tag{2.8.7}$$

where we have used the fact that $\Delta_{T_k} \leq CN^{-1}$ even on the decreasing discretization. We can proceed similarly for $\bar{\epsilon}_2$ from inequality (2.6.45) , to get

$$\begin{aligned}
\bar{\epsilon}_2(k) \leq & C'' \Bigg( (T-T_k)^{-3/2} \Delta^3_{T_{k+1}} + (T-T_k)^{-3/2} \left[ \Delta^{q+1}_{T_{k+1}} + \Delta^2_{T_{k+1}} N^{-[(m-1)\wedge 2q]/2} \right] \\
& + \sum_{i=m-2}^{m+3} (T-T_k)^{-(i-1)/2} \Delta^{(i+1)/2}_{T_{k+1}} \Bigg).
\end{aligned} \tag{2.8.8}$$

Then, (2.8.6) follows by applying Lemma 2.8.1 to (2.8.7) and (2.8.8) □

### *2.8.2* Proof of Corollary 2.2.3

On a first hand, by following the proof of (2.2.2) in Theorem 2.2.1 we get (2.6.6), where the difference between the integral of the $\varphi_i,\ i = 1, \ldots, d$ against the measures in the right hand side are replaced by the distance $d_{\mathcal{F}}$. Since for all $T_k < t < T$ $d_{\mathcal{F}}(\mu_t, \mu_{T_k}) \leq C(t - T_k)$ (resp. $(t - T_k)^{1/2}$) in the case (2.2.7) (resp. (2.2.8)), the result follows from Gronwall's Lemma. This gives the rate of approximation of the law of the forward process.

On a second hand, the backward errors are then obtained by the same arguments already developed in the proof of (2.2.3) in Theorem 2.2.1, using the new forward approximation (2.2.7) (resp. (2.2.8)) instead of (2.2.2) (resp. (2.2.5)) in the proofs of Claims 2.6.5 and 2.6.6.

# ANNEXE B

## Regularization properties of McKean-Vlasov processes

### *B.1* Main results

Consider the following system :

$$\begin{cases} dX_t^x = \sum_{i=0}^d V_i(X_t^x, \mathbb{E}\varphi_i(X_t^x))dB_t^i \\ X_0^x = x \end{cases} \tag{B.1.1}$$

for any $t$ in $[0,T]$, $T>0$ where $(B_t^{1:d})_{t\in[0,T]}$ is a $d$-dimensional Brownian motion defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ and where $B_t^0 = t$ for all $t \in [0,T]$. The vector fields $V_i : (y,w) \in \mathbb{R}^d \times \mathbb{R} \mapsto V_i(y,w)$ are supposed to be bounded and infinitely differentiable with bounded derivatives and the $\varphi_i$, $i = 0, \cdots, d$ are measurable functions from $\mathbb{R}^d$ to $\mathbb{R}$ whose the regularity is precised below.

In this part, we study the differentiability properties of the inhomogeneous semi-group associated to the solution of (B.1.1) and of the solution of (B.1.1) itself with respect to the initial condition.

This study is motivated by the extension of the algorithm presented in Chapter 2 to the class of systems where the $\varphi_i's$ appearing in the coefficients are only Lipschitz continuous and where the matrix[37] $VV^*$ is uniformly elliptic. More precisely, we want to prove that, in this case, the convergence rate of the algorithm presented in Chapter 2 is of order 1 (in term of discretization step).

In order to obtain this result, we expect that the semi-group still regularizes and that for all $t$ in $(0,T]$, the mapping $x \mapsto \mathbb{E}\left[\varphi_i(X_t^x)\right]$ is two times differentiable with bounded derivatives so that we can achieve the analysis done in the previous chapter. Obviously, even if we obtain the usual bounds on the derivatives of the inhomogeneous semi-group w.r.t. the initial condition, this would be not the end of the story. Since in this case the regularization effect is forward, we have to refine the time discretization grid in a suitable way at the beginning of the interval (in addition to the refinement already done at the end) in order to compensate the explosion of the gradient bounds.

As a first step, we investigate the smoothing effect of the inhomogeneous semi-group of the solution of (B.1.1) when the matrix $VV^*$ is uniformly elliptic and when the $\varphi_i's$ appearing in the coefficients are only Lipschitz continuous. In this case, the solution of (B.1.1) admits a transition density. Hence, our investigation is based on a parametrix expansion of this density. Thanks to this representation, we obtain differentiability properties and estimates on the first and second order derivatives of the density in small time. This allows

37. Where $V = [V_1, \cdots, V_d]$ is the $d \times d$ matrix with the vector $V_j$ as $j^{\text{th}}$ column, $j = 1, \cdots, d$.

us to derive usual gradient bounds of the inhomogeneous semi-group. Then, we study the weak differentiabilty (in a $\mathbb{L}_p$, $p \geq 2$ sense) of the solution of (B.1.1) w.r.t. the initial condition.

**Notations and assumptions.** Let

$$\mathcal{L} := V_0 \nabla_y + \frac{1}{2}\mathrm{Tr}[VV^* \nabla_y^2], \tag{B.1.2}$$

where $V = [V_1, \cdots, V_d]$ is the $d \times d$ matrix with the vector $V_j$ as $j^{\text{th}}$ column, $j = 1, \cdots, d$. For any Lipschitz function $\phi$, we denote by $||\phi||_{\text{Lip}}$ its Lipschitz constant : $||\phi||_{\text{Lip}} = \sup_{x \neq y} \big(|\phi(x) - \phi(y)|/|x-y|\big)$.
In the following, $C$, $C'$, $c$, $c'$ etc... denote some positive constants that may change from line to line.

We will consider the two following sets of assumptions :

**(H1):** We say that hypotheses **(H1)** hold if :

(1) Each function $\varphi_i : y \in \mathbb{R}^d \mapsto \varphi_i(y) \in \mathbb{R}$ is supposed to be Lipschitz continuous, $i = 0, \cdots, d$.

(2) There exists $\Lambda > 0$ such that :

$$\forall w \in \mathbb{R},\ \forall y \in \mathbb{R}^d,\ \forall \zeta \in \mathbb{R}^d, \quad \Lambda^{-1}|\zeta|^2 \leq \langle (VV^*)(y,w)\zeta, \zeta\rangle \leq \Lambda|\zeta|^2.$$

**(H2):** We say that hypotheses **(H2)** hold if the mapping $\varphi_i$, $i = 0, \ldots, d$ are bounded and infinitely differentiable with bounded derivatives.

It follows from arguments of [Szn91] that there exists a unique solution $(X_s^{t,x})_{t \leq s \leq T}$ for (B.1.1) under **(H1)** or **(H2)**. For all $t \in [0,T]$, we denote by $\bar{\mu}(t,x) = (\mu_s(t,x))_{t \leq s \leq T}$ the law of the process $(X_s^{t,x})_{t \leq s \leq T}$ (starting from $x$ at time $t$). When $t = 0$, we sometimes write $\bar{\mu}(x)$ instead of $\bar{\mu}(0,x)$ and $(X_s^x)_{0 \leq s \leq T}$ instead of $(X_s^{0,x})_{0 \leq s \leq T}$. More generally, the generic notation $\bar{\nu}$ stands for a family of probability measures $\bar{\nu} = (\nu_t)_{t \leq T}$ on $\mathbb{R}^d$ admitting moments of second order.

**Results.** Our results are summarized by the two following propositions :

**Proposition B.1.1.** *Suppose that hypotheses **(H1)** hold. Let $\phi$ be a Lipschitz continuous function from $\mathbb{R}^d$ to $\mathbb{R}$. Then, for $T$ small enough and for all $t$ in $(0,T]$, the mapping*

$$x \in \mathbb{R}^d \mapsto \mathbb{E}\left[\phi(X_t^x)\right],$$

*where $(X_t^x)_{0 \leq t \leq T}$ is the unique strong solution of (B.1.1), is twice continuously differentiable and for any integer $n$ less than 2,*

$$\frac{\partial^n}{\partial x^n}\mathbb{E}\left[\phi(X_t^x)\right] \leq C\Big(||\phi||_{\text{Lip}}, \{||\varphi_i||_{\text{Lip}},\ i = 0, \ldots, d\}\Big) t^{(-n+1)/2}. \tag{B.1.3}$$

**Proposition B.1.2.** *Suppose that hypotheses **(H2)** hold. Then, for all $t$ in $[0,T]$, the mapping*

$$X_t : x \in \mathbb{R}^d \mapsto X_t^x, \tag{B.1.4}$$

*is bounded and infinitely $\mathbb{L}_p$, $p \geq 2$ - differentiable w.r.t. $x$ with bounded derivatives. Here the boundedness is in $\mathbb{L}_p$, $p \geq 2$.*

## *B.2* Proof of Proposition B.1.1

For all family of probability measures $\bar{\nu} = (\nu_t)_{0\leq t\leq T}$ on $\mathbb{R}^d$ admitting moments of second order, we introduce the process

$$dX_s^{t,y,\bar{\nu}} = \sum_{i=0}^{d} V_i(X_s^{t,y,\bar{\nu}}, \langle \nu_s, \varphi_i\rangle)dB_s^i, \quad X_t^{t,y,\bar{\nu}} = y, \tag{B.2.1}$$

for any $t < s$ in $[0,T]^2$ where "$\langle \nu_s, \varphi_i\rangle$" is the dual notation of "$\int \varphi_i d\nu_s$".

It is clear that for all $\bar{\nu}$, $X^{\cdot,\cdot,\bar{\nu}}$ admits a density and in particular for $\bar{\nu} = \bar{\mu}(t,y)$. For a fixed $\bar{\nu}$ we denote by $p(\bar{\nu};.)$ the transition density of $X^{\cdot,\cdot,\bar{\nu}}$. We have that

$$\mathbb{E}\left[\phi(X_t^{0,y,\bar{\nu}})\right] = \int_{\mathbb{R}} \phi(z)p(\bar{\nu};0,y;t,z)dz, \tag{B.2.2}$$

**First order derivative.** If $x \mapsto \mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right]$ is differentiable, it is then clear that

$$\frac{\partial}{\partial x}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right] = \left[\frac{\partial}{\partial \xi}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(\xi)})\right]\right]_{\xi=x} + \left[\frac{\partial}{\partial y}\mathbb{E}\left[\phi(X_t^{0,y,\bar{\mu}(x)})\right]\right]_{y=x}. \tag{B.2.3}$$

• On a first hand we can deduce from classical arguments (see e.g. [Fri08]) the following Lemma :

**Lemma B.2.1.** *For all family of probability measures $\bar{\nu}$, for all Lipschitz function $\phi$, for all $t$ in $[0,T]$, the mapping $\mathbb{E}\left[\phi(X_t^{0,\cdot,\bar{\nu}})\right] : y \in \mathbb{R}^d \mapsto \mathbb{E}\left[\phi(X_t^{0,y,\bar{\nu}})\right]$ is $n$-times continuously differentiable and there exists a positive constant $C(||\phi||_{\text{Lip}})$ depending on known parameters in **(H1)** and on the Lipschitz constant of $\phi$ such that :*

$$\left|\frac{\partial^n}{\partial y^n}\mathbb{E}\left[\phi(X_t^{0,y,\bar{\nu}})\right]\right| \leq C(||\phi||_{\text{Lip}})t^{(-n+1)/2}.$$

So that, by applying Lemma B.2.1 on the second term in the right hand side of (B.2.3) we have that :

$$\left|\left[\frac{\partial}{\partial y}\mathbb{E}\left[\phi(X_t^{0,y,\bar{\mu}(x)})\right]\right]_{y=x}\right| \leq C(||\phi||_{\text{Lip}}). \tag{B.2.4}$$

• On the other hand by definition :

$$\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(\xi)})\right] = \int_{\mathbb{R}} \phi(z)p(\bar{\mu}(\xi);0,x;t,z)dz, \tag{B.2.5}$$

and we have :

**Lemma B.2.2.** *For all $n \leq 2$, there exist two positive constants $C$ and $c$, depending only on known parameters in **(H1)**, such that for all $t$ in $[0,T]$ and $z, y, \xi$ in $\mathbb{R}^d$*

$$\begin{aligned}
&\left|\frac{\partial^n}{\partial \xi^n} p(\bar{\mu}(t,\xi); t, y; T, z)\right| \\
&\leq C \Bigg\{ (T-t)^{(-n+1)/2} \sum_{i=0}^{d} \sup_{r \in [t,T]} \left| (r-t)^{(n-1)/2} \frac{\partial^n}{\partial \xi^n} \mathbb{E}\left[\varphi_i(X_r^{t,\xi,\bar{\mu}(t,\xi)})\right]\right| \\
&+ \mathbf{1}_{n=2} \sum_{i=0}^{d} \left( \sup_{r \in [t,T]} \left| \frac{\partial}{\partial \xi} \mathbb{E}\left[\varphi_i(X_r^{t,\xi,\bar{\mu}(t,\xi)})\right]\right| \right)^2 \Bigg\} \times \frac{c}{(2\pi(T-t))^{d/2}} \exp\left(-c\frac{|z-y|^2}{(T-t)}\right).
\end{aligned}$$

So that, thanks to Lebesgue differentiation Theorem the first term in the right hand side of (B.2.3) exists and

$$\left[\frac{\partial}{\partial \xi}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(\xi)})\right]\right]_{\xi=x} = \int_{\mathbb{R}} (\phi(z)-\phi(x)) \left[\frac{\partial}{\partial \xi} p(\bar{\mu}(\xi); 0, x; t, z)\right]_{\xi=x} dz.$$

Then, by using Lemma B.2.2 we have :

$$\begin{aligned}
&\left|\left[\frac{\partial}{\partial \xi}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(\xi)})\right]\right]_{\xi=x}\right| \\
&\leq \int_{\mathbb{R}} \Bigg\{ |\phi(z)-\phi(x)| C \sum_{i=0}^{d} \sup_{s\in[0,t]} \left|\frac{\partial}{\partial x}\mathbb{E}[\varphi_i(X_s^{0,x,\bar{\mu}(x)})]\right| \\
&\qquad \times \frac{c}{(2\pi t)^{d/2}} \exp\left(-c\frac{|z-x|^2}{t}\right)\Bigg\} dz.
\end{aligned}$$

So that

$$\begin{aligned}
&\left|\left[\frac{\partial}{\partial \xi}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(\xi)})\right]\right]_{\xi=x}\right| \\
&\leq C||\phi||_{\mathrm{Lip}} t^{1/2} \sum_{i=0}^{d} \sup_{s\in[0,t]} \left|\frac{\partial}{\partial x}\mathbb{E}[\varphi_i(X_s^{0,x,\bar{\mu}(x)})]\right|.
\end{aligned} \tag{B.2.6}$$

• Finally, by plugging (B.2.4) and (B.2.6) in (B.2.3), we obtain that

$$\begin{aligned}
&\left|\frac{\partial}{\partial x}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right]\right| \\
&\leq C(||\phi||_{\mathrm{Lip}}) \left( t^{1/2} \sum_{i=0}^{d} \sup_{s\in[0,t]} \left|\frac{\partial}{\partial x}\mathbb{E}[\varphi_i(X_s^{0,x,\bar{\mu}(x)})]\right| + 1\right).
\end{aligned} \tag{B.2.7}$$

Then, by letting $\phi = \varphi_0$ we deduce that, for $T$ small enough :

$$\left|\frac{\partial}{\partial x}\mathbb{E}\left[\varphi_0(X_t^{0,x,\bar{\mu}(x)})\right]\right|$$
$$\leq C(||\varphi_0||_{\mathrm{Lip}})\left[1+t^{1/2}\sum_{i=1}^{d}\sup_{s\in[0,t]}\left|\frac{\partial}{\partial x}\mathbb{E}[\varphi_i(X_s^{0,x,\bar{\mu}(x)})]\right|\right].$$

We can plug this estimate in (B.2.7), and iterate this argument : letting $\phi=\varphi_1$ etc... Finally, we deduce that for any Lipschitz function $\phi$, for all $t\in[0,T]$ with $T$ small enough, the mapping $x\mapsto\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right]$ is continuously differentiable and :

$$\left|\frac{\partial}{\partial x}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right]\right|\leq C(||\phi||_{\mathrm{Lip}},||\varphi_{0:d}||_{\mathrm{Lip}}). \tag{B.2.8}$$

This concludes the proof for the first order derivative. □

**Second order derivative.** Now, we come back to (B.2.2). If $x\mapsto\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right]$ is two times differentiable, then :

$$\begin{aligned}\frac{\partial^2}{\partial x^2}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right] &= \left[\frac{\partial^2}{\partial \xi^2}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(\xi)})\right]\right]_{\xi=x}+\left[\frac{\partial^2}{\partial y^2}\mathbb{E}\left[\phi(X_t^{0,y,\bar{\mu}(x)})\right]\right]_{y=x}\\ &\quad+2\left[\frac{\partial^2}{\partial\xi\partial y}\mathbb{E}\left[\phi(X_t^{0,y,\bar{\mu}(\xi)})\right]\right]_{(\xi,y)=(x,x)}.\end{aligned} \tag{B.2.9}$$

From Lemma B.2.2, and (B.2.5), Lemma B.2.1 and Lebesgue differentiation Theorem, the two first terms in the right hand side exist and they are bounded as follows :

$$\left|\left[\frac{\partial^2}{\partial y^2}\mathbb{E}\left[\phi(X_t^{0,y,\bar{\mu}(x)})\right]\right]_{y=x}\right|\leq t^{-1/2}C(||\phi||_{\mathrm{Lip}}). \tag{B.2.10}$$

and

$$\begin{aligned}\left|\left[\frac{\partial^2}{\partial \xi^2}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(\xi)})\right]\right]_{\xi=x}\right| &\leq C\Bigg\{\sum_{i=0}^{d}\sup_{r\in[0,t]}\left|r^{1/2}\frac{\partial^2}{\partial x^2}\mathbb{E}\left[\varphi_i(X_r^{0,x,\bar{\mu}(x)})\right]\right|\\ &\quad+t^{1/2}\sum_{i=0}^{d}\left(\sup_{r\in[0,t]}\left|\frac{\partial}{\partial x}\mathbb{E}\left[\varphi_i(X_r^{0,x,\bar{\mu}(x)})\right]\right|\right)^2\Bigg\}.\end{aligned}$$

Let us deal with the third term in the right hand side of (B.2.9). We have :

**Lemma B.2.3.** *There exist two positive constants $C$ and $c$, depending only on known parameters in* ***(H1)****, such that for all $t$ in $[0,T]$ and $z,y,\xi$ in $\mathbb{R}^d$*

$$\left|\frac{\partial^2}{\partial\xi\partial y}p(\bar{\mu}(t,\xi);t,y;T,z)\right|$$
$$\leq C(T-t)^{-1/2}\left\{\sum_{i=0}^{d}\sup_{r\in[t,T]}\frac{\partial}{\partial\xi}\mathbb{E}\left[\varphi_i(X_r^{t,\xi,\bar{\mu}(t,\xi)})\right]\right\}\frac{c}{(2\pi(T-t))^{d/2}}\exp\left(-c\frac{|z-y|^2}{(T-t)}\right).$$

Hence, thanks to Lebesgue differentiation Theorem :

$$\left[\frac{\partial^2}{\partial\xi\partial y}\mathbb{E}\left[\phi(X_t^{0,y,\bar{\mu}(\xi)})\right]\right]_{(\xi,y)=(x,x)}=\int_{\mathbb{R}}(\phi(z)-\phi(x))\left[\frac{\partial^2}{\partial\xi\partial y}p(\bar{\mu}(\xi);0,x;t,z)\right]_{(\xi,y)=(x,x)}dz,$$

and we have that :

$$\left|\left[\frac{\partial^2}{\partial\xi\partial y}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(\xi)})\right]\right]_{(\xi,y)=(x,x)}\right| \tag{B.2.11}$$
$$\leq C||\phi||_{\mathrm{Lip}}\sum_{i=0}^{d}\sup_{s\in[0,t]}\left|\frac{\partial}{\partial x}\mathbb{E}\left[\varphi_i(X_s^{0,x,\bar{\mu}(x)})\right]\right|.$$

- Finally, by plugging (B.2.10), (B.2.11) and (B.2.11) in (B.2.9), we obtain that

$$\left|\frac{\partial^2}{\partial x^2}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right]\right| \tag{B.2.12}$$
$$\leq C(||\phi||_{\mathrm{Lip}})\Bigg\{\sum_{i=0}^{d}\sup_{s\in[0,t]}s^{1/2}\left|\frac{\partial}{\partial x}\mathbb{E}[\varphi(X_s^{0,x,\bar{\mu}(x)})]\right|+t^{-1/2}$$
$$+\sum_{i=0}^{d}\left(\sup_{s\in[0,t]}\left|\frac{\partial}{\partial x}\mathbb{E}\left[\varphi_i(X_s^{0,x,\bar{\mu}(x)})\right]\right|\right)^2+\sum_{i=0}^{d}\sup_{s\in[0,t]}\left|\frac{\partial}{\partial x}\mathbb{E}\left[\varphi_i(X_s^{0,x,\bar{\mu}(x)})\right]\right|\Bigg\}.$$

Then, by plugging estimate (B.2.8) in (B.2.12), multiplying both sides by $t^{1/2}$ and letting $\phi=\varphi_0$ we obtain that :

$$\left|\frac{\partial^2}{\partial x^2}\mathbb{E}\left[\varphi_0(X_t^{0,x,\bar{\mu}(x)})\right]\right|$$
$$\leq C(||\varphi_0||_{\mathrm{Lip}})\left[1+t^{1/2}\sum_{i=1}^{d}\sup_{s\in[0,t]}\left|s^{1/2}\frac{\partial^2}{\partial x^2}\mathbb{E}[\varphi_i(X_s^{0,x,\bar{\mu}(x)})]\right|\right],\ \text{for } T \text{ small enough.}$$

We can plug this estimate in (B.2.12) and iterate this argument : multiplying both sides by $t^{1/2}$, letting $\phi=\varphi_1$ etc... Finally, we deduce that for any Lipschitz function $\phi$, for all $t\in[0,T]$ with $T$ small enough, the mapping $x\mapsto\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right]$ is two times continuously differentiable and satisfies :

$$\left|\frac{\partial^2}{\partial x^2}\mathbb{E}\left[\phi(X_t^{0,x,\bar{\mu}(x)})\right]\right|\leq C(||\phi||_{\mathrm{Lip}},||\varphi_{0:d}||_{\mathrm{Lip}})t^{-1/2}. \tag{B.2.13}$$

### *B.2.1* Proof of Lemma B.2.2

Here, we show that the solution of (B.1.1) admits a transition density $p$ such that for all $t\in[0,T]$ and $y,z\in\mathbb{R}^d$, the mapping $x\mapsto p(\bar{\mu}(t,x);t,y;T,z)$ is twice continuously differentiable and for all integer $n$ less than 2 :

$$\left|\frac{\partial^n}{\partial x^n}p(\bar{\mu}(t,x);t,y;T,z)\right| \leq C\big\{(T-t)^{(-n+1)/2}F_1^n(t,t,T)$$
$$+\mathbf{1}_{n=2}F_2^1(t,t,T)\big\}\frac{1}{(2\pi(T-t))^{d/2}}\exp\left(-c\frac{|z-y|^2}{(T-t)}\right),$$

where for all $s$ in $(t,T]$ :

$$F_k^l(t,s,T) = \sum_{i=1}^d \left( \sup_{r\in[s,T]} \left| (r-t)^{(l-1)/2} \frac{\partial^l}{\partial x^l} h_i(t,x;r) \right| \right)^k \quad \text{and } h_i(t,x;r) = \langle \mu_r(t,x), \varphi_i \rangle, \tag{B.2.14}$$

for $k,l = 1,2$.

The proof is based on a parametrix expansion of the transition density of the solution of (B.1.1). Given a family of probability measures $\bar{\nu} = (\nu)_{0\le t\le T}$ on $\mathbb{R}^d$ we recall that the solution of

$$dX_s^{t,y,\bar{\nu}} = \sum_{i=0}^d V_i(X_s^{t,y,\bar{\nu}}, \langle \nu_s, \varphi_i \rangle) dB_s^i, \quad X_t^{t,y,\bar{\nu}} = y,$$

admits a density $p(\bar{\nu};.)$ for all $\bar{\nu}$ and that this density solves the Fokker-Planck equation :

$$\begin{aligned}
&\frac{\partial}{\partial t} p(\bar{\nu};t,y;T,z) + \mathcal{L}_y^{\bar{\nu}} p(\bar{\nu};t,y;T,z) = 0, \quad 0\le t\le T,\ y,z\in\mathbb{R}^d,\\
&p(\bar{\nu};T,y;T,z) = \delta_z(y), \quad y,z\in\mathbb{R}^d,
\end{aligned}$$

where $\mathcal{L}_y^{\bar{\nu}}$ is the generator of the diffusion process $X^{\cdot,y,\bar{\nu}}$.

*(i) The frozen system.* For all $\zeta\in\mathbb{R}^d$ and for all family of probability measures $\bar{\nu}$, we define the frozen process $\tilde{X}^{\cdot,\cdot,\bar{\nu},\zeta}$ as the solution [38] of

$$d\tilde{X}_s^{t,y,\bar{\nu},\zeta} = \sum_{i=0}^d V_i(\zeta, \langle \nu_s, \varphi_i \rangle) dB_s^i, \quad \tilde{X}_t^{t,y,\bar{\nu},\zeta} = y.$$

This process admits a transition density $\tilde{p}^\zeta(\bar{\nu};\cdot)$ which solves :

$$\begin{aligned}
&\frac{\partial}{\partial t} \tilde{p}^\zeta(\bar{\nu};t,y;T,z) + \tilde{\mathcal{L}}_y^{\bar{\nu},\zeta} \tilde{p}^\zeta(\bar{\nu};t,y;T,z) = 0, \quad 0\le t\le T,\ y,z\in\mathbb{R}^d,\\
&\tilde{p}^\zeta(\bar{\nu};T,y;T,z) = \delta_z(y) \quad y,z\in\mathbb{R}^d.
\end{aligned} \tag{B.2.15}$$

where $\tilde{\mathcal{L}}_y^{\bar{\nu},\zeta}$ is the generator of $\tilde{X}^{\cdot,y,\bar{\nu},\zeta}$. Moreover, this density is Gaussian and for all $0\le t\le T,\ y,z\in\mathbb{R}^d$ it is given by :

$$\tilde{p}^\zeta(\bar{\nu};t,y;T,z) = \frac{1}{(2\pi)^{d/2}} \left[ \det\left(\tilde{\Sigma}_{t,T}^\zeta(\bar{\nu})\right)\right]^{-1/2} \exp\left( -|(\tilde{\Sigma}_{t,T}^\zeta(\bar{\nu}))^{-1/2}(z - \tilde{m}_{t,T}^\zeta(\bar{\nu},y))|^2 \right), \tag{B.2.16}$$

where

$$\tilde{m}_{t,T}^\zeta(\bar{\nu},y) = y + \int_t^T V_0(\zeta, \langle \nu_s, \varphi_0 \rangle) ds,$$

and where $\tilde{\Sigma}_{t,T}^\zeta(\bar{\nu})$ is given by :

$$\tilde{\Sigma}_{t,T}^\zeta(\bar{\nu}) = \int_t^T VV^*(\zeta, \langle \nu_s, \varphi_{1:d} \rangle) ds.$$

38. Here the subscript $\zeta$ stands for the dependence on the freezing point.

Here, the notation "1 : $d$" means "from 1 to $d$".

It is clear that there exist two positive constants $C$ and $c$, depending only on known parameters in **(H1)** and a Gaussian type function $\hat{p}_c$ defined by :

$$\hat{p}_c(t, y, ; T, z) = \frac{1}{(2\pi(T-t))^{d/2}} \exp\left(-c\frac{|z-y|^2}{(T-t)}\right),$$

such that for all family of probability measures $\bar{\nu}$, for all freezing point $\zeta$ in $\mathbb{R}^d$ :

$$\tilde{p}^{\zeta}(\bar{\nu}; t, y; T, z) \leq C\hat{p}_c(t, y; T, z), \tag{B.2.17}$$

for all $0 \leq t \leq T$ and $y, z \in \mathbb{R}^d$. Moreover, for any $n$ in $\mathbb{N}^*$ it is clear that there exists a positive constant $C$ depending only on known parameters in **(H1)** such that :

$$\left|\frac{\partial^n}{\partial y^n}\tilde{p}^{\zeta}(\bar{\nu}; t, y; T, z)\right| \leq C(T-t)^{-n/2}\hat{p}_c(t, y; T, z), \tag{B.2.18}$$

for all $0 \leq t \leq T$ and $y, z \in \mathbb{R}^d$.

*(ii) The parametrix expansion.* Let $z$ in $\mathbb{R}^d$, for all $\zeta \in \mathbb{R}^d$, for all $\bar{\nu}$, the transition density $\tilde{p}^{\zeta}(\bar{\nu}; \cdot, \cdot; T, z)$ satisfies the Fokker-Planck equation (B.2.15) which can be written as :

$$\frac{\partial}{\partial t}\tilde{p}^{\zeta}(\bar{\nu}; t, y; T, z) + \mathcal{L}^{\bar{\nu}}\tilde{p}^{\zeta}(\bar{\nu}; t, y; T, z) = \left(\mathcal{L}^{\bar{\nu}}_y - \tilde{\mathcal{L}}^{\bar{\nu},\zeta}_y\right)\tilde{p}^{\zeta}(\bar{\nu}; t, y; T, z), \quad 0 \leq t \leq T, y, z \in \mathbb{R}^d$$
$$\tilde{p}^{\zeta}(\bar{\nu}; T, y; T, z) = \delta_z(y), \quad y, z \in \mathbb{R}^d.$$

Note that $p(\bar{\nu}; ., .; T, z)$ is a fundamental solution of this PDE. Therefore, by choosing $\zeta = z$ (the freezing point to be the arrival point), $\tilde{p}^z(\bar{\nu}; ., .; T, z)$ writes :

$$p(\bar{\nu}; t, y; T, z) - \tilde{p}^z(\bar{\nu}; t, y; T, z) = -\int_t^T \int_{\mathbb{R}^d} \left(\mathcal{L}^{\bar{\nu}}_u - \tilde{\mathcal{L}}^{\bar{\nu},\zeta}_u\right)\tilde{p}^z(\bar{\nu}; s, u; T, z)p(\bar{\nu}; t, y; s, u)duds. \tag{B.2.19}$$

By iterating $N$ times this procedure, we obtain that :

$$\begin{aligned} p(\bar{\nu}; t, y; T, z) &= \tilde{p}^z(\bar{\nu}; t, y; T, z) + \sum_{k=1}^{N}\int_t^T \int_{\mathbb{R}^d} H^{\otimes k}(\bar{\nu}; s, u; T, z)\tilde{p}^u(\bar{\nu}; t, y; s, u)duds \\ &\quad + \int_t^T \int_{\mathbb{R}^d} H^{\otimes N+1}(\bar{\nu}; s, u; T, z)p(\bar{\nu}; t, y; s, u)duds, \end{aligned} \tag{B.2.20}$$

where

$$H(\bar{\nu}; t, y; T, z) = \left(\tilde{\mathcal{L}}^{\bar{\nu},\zeta}_y - \mathcal{L}^{\bar{\nu}}_y\right)\tilde{p}^z(\bar{\nu}; t, y; T, z), \tag{B.2.21}$$

and where $H^{\otimes k}$ is recursively defined by :

$$H^{\otimes k+1}(\bar{\nu}; t, y; T, z) = \int_t^T \int_{\mathbb{R}^d} H^{\otimes k}(\bar{\nu}; s, u; T, z)H(\bar{\nu}; t, y; s, u)duds. \tag{B.2.22}$$

In order to obtain a parametrix expansion of $p(\bar{\nu}; \cdot)$, depending only on known quantities (i.e. on the regularization kernel $H$ defined by (B.2.21) and on the transition density of the frozen process $\tilde{p}(\bar{\nu}; \cdot)$) the idea consists in letting $N$ tends to infinity.

To this aim, we need a "good" estimate on the approximation error. Here these controls are :

(i) The Gaussian estimate on $\tilde{p}^{\zeta}(\bar{\nu};\cdot)$ given by (B.2.17)

(ii) The following Claim :

**Claim B.2.4.** *For all $k>0$, there exists a positive constant $C_k$ given by :*

$$C_k = \frac{C^k 4^k}{\sqrt{(k-1)!}} := \frac{\gamma^k}{\sqrt{(k-1)!}},$$

*such that*

$$|H^{\otimes k}(\bar{\nu}; s, y; T, z)| \leq C_k (T-s)^{k/2-1} \hat{p}_c(s, y; T, z).$$

On the one hand, the first term in the right hand side of (B.2.20) is controlled by a convolution of two Gaussian functions which is still Gaussian. On the other hand, the $N^{\text{th}}$ convolution of the kernel $H$ tends uniformly to 0 as $N$ tends to infinity (recall that $T$ is small) and since $p(\bar{\nu};\cdot)$ is a density, we deduce that the second term in the right hand side of (B.2.20) tends to 0. Therefore, the density $p(\bar{\nu};\cdot)$ writes :

$$\begin{aligned} p(\bar{\nu}; t, y; T, z) &= \tilde{p}^{z}(\bar{\nu}; t, y; T, z) \\ &\quad + \sum_{k\geq 1} \int_t^T \int_{\mathbb{R}^d} H^{\otimes k}(\bar{\nu}; s, u; T, z) \tilde{p}^{u}(\bar{\nu}; t, y; s, u) du ds. \end{aligned} \tag{B.2.23}$$

*(iii) Conservation of the Gaussian decay when $\bar{\nu} = \bar{\mu}(t,x)$.* We now study the case when the family of probability measures depends on parameters $(t,x)$ in $[0,T]\times\mathbb{R}^d$. In particular, we take $\bar{\nu} = \bar{\mu}(t,x)$. We want to show that the derivative of the transition density (B.2.23) w.r.t. the space parameter is bounded from above by a Gaussian type function.

Let us first remark that, formally (if Lebesgue differentiation Theorem applies), the derivative of $p$ w.r.t. $x$ writes :

$$\begin{aligned} &\frac{\partial}{\partial x} p(\bar{\mu}(t,x); t, y; T, z) \\ &= \frac{\partial}{\partial x} \tilde{p}^{z}(\bar{\mu}(t,x); t, y; T, z) + \sum_{k\geq 1} \int_t^T \int_{\mathbb{R}^d} \frac{\partial H^{\otimes k}}{\partial x}(\bar{\mu}(t,x); s, u; T, z) \tilde{p}^{u}(\bar{\mu}(t,x); t, y; s, u) du ds \\ &\quad + \sum_{k\geq 1} \int_t^T \int_{\mathbb{R}^d} H^{\otimes k}(\bar{\mu}(t,x); s, u; T, z) \frac{\partial \tilde{p}^{u}}{\partial x}(\bar{\mu}(t,x); t, y; s, u) du ds. \end{aligned} \tag{B.2.24}$$

At the end of this Section, we prove that the derivatives that appear in the integrand of (B.2.24) when differentiating the function $x \mapsto p(\bar{\mu}(\cdot,x),\cdot)$ are still controlled by Gaussian type functions :

**Claim B.2.5.** *There exist two positive constants $C$ and $c$ such that, for all $(t,y,z)$ in $[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$ :*

$$\left|\frac{\partial^n}{\partial x^n} \tilde{p}(\bar{\mu}(t,x); t, y; T, z)\right| \leq C\big\{(T-t)^{(-n+1)/2} F_1^n(t,t,T) + \mathbf{1}_{n=2} F_2^1(t,t,T)\big\} \hat{p}_c(s, y; T, z).$$

*and for all $s \in (t,T]$ :*

$$\left|\frac{\partial^n}{\partial x^n} \tilde{p}(\bar{\mu}(t,x); s, y; T, z)\right| \leq C\big\{(s-t)^{(-n+1)/2} F_1^n(t,s,T) + \mathbf{1}_{n=2} F_2^1(t,s,T)\big\} \hat{p}_c(s, y; T, z)$$

*for all $n \leq 2$ and where $\{F_k^n(t,s,T),\ k=1,2,\ n=1,2\}$ is defined by (B.2.14).*

**Claim B.2.6.** *For all $k > 0$, there exists a positive constants $\tilde{C}_k$ satisfying :*

$$\tilde{C}_k \leq (k-1)\frac{\gamma^{k-1}}{\sqrt{(k-1)!}} + \frac{C^k 4^k}{\sqrt{(k-1)!}}$$

*such that*

$$\left|\frac{\partial^n}{\partial x^n} H^{\otimes k}(\bar{\mu}(t,x); s, y; T, z)\right| \leq \Big\{(s-t)^{(-n+1)/2} F_1^n(t,s,T) + \mathbf{1}_{n=2} F_2^1(t,s,T)\Big\} \times \tilde{C}_k (T-s)^{k/2-1} \hat{p}_c(s,y;T,z),$$

*for all $n \leq 2$.*

By plugging the estimates of Claims B.2.4, B.2.5 and B.2.6 in (B.2.24) and using Gaussian convolution, we deduce from Lebesgue differentiation Theorem that there exist two positive constants $C$ and $c$, depending only on known parameters in **(H1)**, such that :

$$\left|\frac{\partial^n}{\partial x^n} p(\bar{\mu}(t,x); t, y; T, z)\right| \leq C\big\{(T-t)^{(-n+1)/2} F_1^n(t,t,T) + \mathbf{1}_{n=2} F_2^1(t,t,T)\big\}\hat{p}_c(t,y;T,z)$$

This concludes the proof of Lemma B.2.2. □

### *B.2.2* Proof of Claims B.2.4, B.2.5 and B.2.6

***B.2.2.1 Proof of Claim B.2.4*** By using classical parametrix arguments (see Chapter 1 of [Fri64]) and by the definition (B.2.21) of $H$, we deduce that there exist two positive constants $C$ and $c$ such that for all $t \leq s < \tau \in [0,T]^3$ and $x, y, z \in \mathbb{R}^d$ :

$$|H(\bar{\mu}(t,x); s, y; \tau, z)| \leq C(\tau - s)^{-1/2} \hat{p}_c(s, y; \tau, z). \tag{B.2.25}$$

Let $x$ in $\mathbb{R}^d$ and suppose now as an induction hypothesis that for all $t \leq s \in [0,T)^2$ and $y, z \in \mathbb{R}^d$.

$$|H^{\otimes k}(\bar{\mu}(t,x); s, y; T, z)| \leq C_k (T-s)^{k/2-1} \hat{p}_c(s, y; T, z), \tag{B.2.26}$$

where

$$C_k = \frac{C^k \sqrt{2}^k}{\sqrt{(k-1)!}} := \frac{\gamma^k}{\sqrt{(k-1)!}}. \tag{B.2.27}$$

Let $t \leq s \in [0,T]^2$ and $y, z \in \mathbb{R}^d$. We recall that for all $k \geq 0$, $H^{\otimes k+1}$ is defined by

$$H^{\otimes k+1}(\bar{\mu}(t,x); s, y; T, z) = \int_s^T \int_{\mathbb{R}^d} H^{\otimes k}(\bar{\mu}(t,x); r, u; T, z) H(\bar{\mu}(t,x); s, y; r, u) du dr. \tag{B.2.28}$$

Hence, by plugging (B.2.25) and (B.2.26) in (B.2.28), we obtain that :

$$|H^{\otimes k+1}(\bar{\mu}(t,x); s, y; T, z)| \leq C_k C \int_s^T (T-r)^{k/2-1}(r-t)^{-1/2} dr \hat{p}_c(s, y; T, z),$$

and by the change of variable $r = (T-s)u + s$ we have :

$$\begin{aligned} &|H^{\otimes k+1}(\bar{\mu}(t,x); s, y; T, z)| \\ &\leq C C_k (T-s)^{(k-1)/2} \int_0^1 (1-u)^{k/2-1} u^{-1/2} du \hat{p}_c(s, y; T, z). \end{aligned} \tag{B.2.29}$$

Note that for all $0<\epsilon<1$, we have :

$$\begin{aligned}
\int_0^1 (1-u)^{k/2-1}u^{-1/2}du &= \int_0^1 (1-u)^{-1/2}u^{k/2-1}du \qquad \text{(B.2.30)}\\
&= \int_0^{1-\epsilon} (1-u)^{-1/2}u^{k/2-1}du \\
&\quad + \int_{1-\epsilon}^1 (1-u)^{-1/2}u^{k/2-1}du \\
&\leq \frac{1}{\sqrt{\epsilon}}\int_0^{1-\epsilon} u^{k/2-1}du + \int_{1-\epsilon}^1 (1-u)^{-1/2}du \\
&= \frac{1}{\sqrt{\epsilon}}\frac{2}{k} + 2\sqrt{\epsilon}.
\end{aligned}$$

So that, by letting $\epsilon = 1/k$ we have

$$\int_0^1 (1-u)^{k/2-1}u^{-1/2}du \leq \frac{4}{\sqrt{k}} \qquad \text{(B.2.31)}$$

Therefore, by plugging (B.2.31) in (B.2.29) we obtain that

$$|H^{\otimes k+1}(\bar{\mu}(t,x);s,y;T,z)| \leq (T-s)^{(k-1)/2}CC_k\frac{4}{\sqrt{k}}\frac{1}{\sqrt{2\pi(T-s)}^d}\exp\left(-c\frac{|z-y|^2}{2(T-s)}\right),$$

and so :

$$C_{k+1} = CC_k\frac{4}{\sqrt{k}} = \frac{C^{k+1}4^{k+1}}{\sqrt{k!}} = \frac{\gamma^{k+1}}{\sqrt{k!}}.$$

□

***B.2.2.2 Proof of Claim B.2.5*** For the reader convenience, we make the proof for $d=n=1$. For all $t<s\in[0,T)^2$ and $y,z\in\mathbb{R}^d$, it is clear that the mapping $x\in\mathbb{R}^d\mapsto \tilde{p}^{\zeta}(\bar{\mu}(t,x);s,y;T,z)$ is continuously differentiable. Let $t<s\in[0,T)^2$ and $y,z\in\mathbb{R}^d$, by setting $a:=VV^*$, we have from definition (B.2.16) of $\tilde{p}$ :

$$\begin{aligned}
&\frac{\partial}{\partial x}\tilde{p}^z(\bar{\mu}(t,x);s,y;T,z) \qquad \text{(B.2.32)}\\
&= -\frac{1}{2\sqrt{2\pi}}\left\{\frac{\int_s^T[\partial a/\partial w](z,h_1(t,x;r))[\partial/\partial x]h_1(t,x;r)dr}{\left(\int_s^T a(z,h_1(t,x;r))dr\right)^{3/2}}\right.\\
&\quad -\frac{\int_s^T[\partial a/\partial w](z,h_1(t,x;r))[\partial/\partial x]h_1(t,x;r)dr\left|z-m^z_{s,T}(\bar{\mu}(t,x),y)\right|^2}{2\left(\int_s^T a(z,h_1(r,x))dr\right)^{5/2}}\\
&\quad \left.+2\frac{\int_s^T[\partial V_0/\partial w](z,h_0(t,x;r))[\partial/\partial x]h_0(t,x;r)dr(z-m^z_{s,T}(\bar{\mu}(t,x),y))}{\left(\int_s^T a(z,h_1(t,x;r))dr\right)^{3/2}}\right\}\\
&\quad \times\exp\left(-\frac{1}{2}\frac{\left|z-m^z_{s,T}(\bar{\mu}(t,x),y)\right|^2}{\int_s^T a(z,h_1(t,x;r))dr}\right)
\end{aligned}$$

We can bound the first term in the right hand side of (B.2.32) in the following way :

$$\begin{aligned}
&\left|\int_s^T [\partial a/\partial w](z,h_1(t,x;r))[\partial/\partial x]h_1(t,x;r)dr\right| \\
&\leq \int_s^T |[\partial a/\partial w](z,h_1(t,x;r))|\,|[\partial/\partial x]h_1(t,x;r)|\,dr \\
&\leq \int_s^T |[\partial a/\partial w](z,h_1(t,x;r))| \sup_{u\in[t,T]} |[\partial/\partial x]h_1(t,x;u)|\,dr \\
&\leq C(T-s) \sup_{r\in[t,T]} [|[\partial/\partial x]h_1(t,x;r)|]\,.
\end{aligned}$$

By using the fact that, for all $\epsilon > 0$ there exists a positive constant $C$ such that :

$$\Lambda\frac{|z-m^z_{s,T}(\bar{\mu}(t,x),y)|}{(T-s)^{1/2}}\exp\left(-\frac{\Lambda\epsilon}{2}\frac{\left|z-m^z_{s,T}(\bar{\mu}(t,x),y)\right|^2}{(T-s)}\right)\leq C,$$

and following the same arguments as above, we can bound the second and the third terms in the right hand side of (B.2.32). We deduce that there exists a positive constant $C'$ such that :

$$\left|\frac{\partial}{\partial x}\tilde{p}^z(\bar{\mu}(t,x);s,y;T,z)\right| \quad\leq\quad F^1_1(t,s,T)C'\hat{p}_c(s,y;T,z),$$

which proves the second assertion of the Claim.

When differentiating a second time, we obtain a bound depending on the product of the derivatives "$([\partial/\partial x]h_k)^2$" and the second order derivative $[\partial^2/\partial x^2]h_k$, $k=0,1$. We bound this last term as follows :

$$\left|[\partial^2/\partial x^2]h_k(t,x;r)dr\right| \leq (r-t)^{-1/2}\sup_{r\in[t,T]}(r-t)^{-1/2}\left|[\partial^2/\partial x^2]h_k(t,x;r)\right|dr,$$

$k=0,1$, and following the arguments above we obtain the corresponding bound in the claim. $\square$

***B.2.2.3 Proof of Claim B.2.6*** Again we make the proof for $d=n=1$. Let $t<s\in[0,T]^2$ and $y,z\in\mathbb{R}^d$. By definition (B.2.21) of $H$, we have :

$$\begin{aligned}
&\frac{\partial}{\partial x}H(\bar{\mu}(t,x);s,y;T,z) \\
&=\Bigg[\left(\frac{\partial V_0}{\partial w}(z,h_0(t,x;s))-\frac{\partial V_0}{\partial w}(y,h_0(t,x;s))\right)[\partial/\partial x]h_0(t,x;s)\frac{\partial}{\partial y} \\
&\quad+\frac{1}{2}\left(\frac{\partial a}{\partial w}(z,h_1(t,x;s))-\frac{\partial a}{\partial w}(y,h_1(t,x;s))\right)[\partial/\partial x]h_1(t,x;s)\frac{\partial^2}{\partial y^2}\Bigg]\tilde{p}^z(\bar{\mu}(t,x);s,y;T,z) \\
&\quad+\left(V_0(z,h(t,x;s))-V_0(y,h_0(t,x;s))\right)\frac{\partial}{\partial x}\frac{\partial\tilde{p}^z}{\partial y}(\bar{\mu}(t,x);s,y;T,z) \\
&\quad+\frac{1}{2}\left(a(z,h_1(t,x;s))-a(y,h_1(t,x;s))\right)\frac{\partial}{\partial x}\frac{\partial^2\tilde{p}^z}{\partial y^2}(\bar{\mu}(t,x);s,y;T,z).
\end{aligned}$$

By using the second estimate in Claim B.2.5 and by noticing that, from the proof of Claim B.2.5 we have :

$$\left|\frac{\partial^{n+1}}{\partial x \partial y^n} p(\bar{\mu}(t,x); s, y; T, z)\right| \leq CF_1^1(t,s,T)\tilde{C}_1(T-s)^{-n/2}\hat{p}_c(s,y;T,z),$$

for any $n \leq 2$, we get :

$$\left|\frac{\partial H}{\partial x}(\bar{\mu}(t,x); s, y; T, z)\right| \leq CF_1^1(t,s,T)\tilde{C}_1(T-s)^{-1/2}\hat{p}_c(s,y;T,z). \quad \text{(B.2.33)}$$

Assume now as an induction hypothesis that for all $(r,u) \in [t,T] \times \mathbb{R}^d$ :

$$\left|\frac{\partial}{\partial x} H^{\otimes k}(\bar{\mu}(t,x); r, u; T, z)\right| \leq CF_1^1(t,r,T)\tilde{C}_k(T-r)^{k/2-1}\hat{p}_c(r,u;T,z), \quad \text{(B.2.34)}$$

where

$$\tilde{C}_k \leq (k-1)\frac{\gamma^{k-1}}{\sqrt{(k-1)!}} + \frac{C^k 4^k}{\sqrt{(k-1)!}},$$

for some positive $\gamma$.

We have

$$\begin{aligned}
&\frac{\partial}{\partial x} H^{\otimes k+1}(\bar{\mu}(t,x); s, y; T, z) \\
&= \int_s^T \int_{\mathbb{R}^d} \frac{\partial}{\partial x} H^{\otimes k}(\bar{\mu}(t,x); r, u; T, z) H(\bar{\mu}(t,x); s, y; r, u) du dr \quad \text{(B.2.35)} \\
&\quad + \int_s^T \int_{\mathbb{R}^d} H^{\otimes k}(\bar{\mu}(t,x); s, u; T, z) \frac{\partial}{\partial x} H(\bar{\mu}(t,x); s, y; r, u) du dr. \quad \text{(B.2.36)}
\end{aligned}$$

We can bound the first term in the right hand side by using the induction hypothesis (B.2.34) and estimate (B.2.25) :

$$\begin{aligned}
|\text{(B.2.35)}| &\leq C\tilde{C}_k \int_s^T (T-r)^{k/2-1}(r-s)^{-1/2}F_1^1(t,r,T)dr\hat{p}_c(s,y;T,z) \\
&\leq C\tilde{C}_k F_1^1(t,s,T) \int_s^T (T-r)^{k/2-1}(r-s)^{-1/2}dr \\
&\quad \times \hat{p}_c(s,y;T,z),
\end{aligned}$$

since for all $r \geq s$ in $(t,T]$ we have that $F_1^1(t,r,T) \leq F_1^1(t,s,T)$. By the change of variable $r = (T-s)u + s$ we have

$$\begin{aligned}
|\text{(B.2.35)}| &\leq F_1^1(t,s,T) \\
&\quad \times C\tilde{C}_k(T-s)^{(k-1)/2}\left(\int_0^1 (1-u)^{k/2-1}u^{-1/2}du\right)\hat{p}_c(s,y;T,z).
\end{aligned}$$

Hence, by (B.2.30),

$$|\text{(B.2.35)}| \leq F_1^1(t,s,T)C\tilde{C}_k(T-s)^{(k-1)/2}\frac{4}{\sqrt{k}}\hat{p}_c(s,y;T,z). \quad \text{(B.2.37)}$$

Now, we bound the term (B.2.36). From (B.2.26) and (B.2.33) we have :

$$\begin{aligned}|(\text{B.2.36})| &\leq C_k\tilde{C}\int_s^T (T-r)^{k/2-1}(r-s)^{-1/2}F_1^1(t,r,T)dr\hat{p}_c(s,y;T,z)\\ &\leq F_1^1(t,s,T)C_k\tilde{C}\frac{4}{\sqrt{k}}(T-s)^{k/2-1}\hat{p}_c(s,y;T,z), \qquad (\text{B.2.38})\end{aligned}$$

where $C_k$ is defined by (B.2.27). From (B.2.37) and (B.2.38), we deduce that :

$$\left|\frac{\partial}{\partial x}H^{\otimes k+1}(\bar{\mu}(t,x);r,u;T,z)\right| \leq CF_1^1(t,r,T)\tilde{C}_{k+1}(T-r)^{(k+1)/2-1}\hat{p}_c(r,u;T,z),$$

where

$$\tilde{C}_{k+1} = C_k\tilde{C}\frac{4}{\sqrt{k}} + \tilde{C}_kC\frac{4}{\sqrt{k}}, \qquad (\text{B.2.39})$$

and the induction is true since

$$\tilde{C}_{k+1} \leq k\frac{\gamma^k}{\sqrt{k!}} + \frac{C^{k+1}4^{k+1}}{\sqrt{k!}},$$

for some positive real $\gamma \geq 4C$. Indeed, from (B.2.39) we have that :

$$\tilde{C}_{k+1} \leq \frac{\gamma^k\tilde{C}}{\sqrt{k!}} + \tilde{C}_kC\frac{4}{\sqrt{k}},$$

where we replaced $C_k$ by its value in Claim B.2.4 so that

$$\frac{\tilde{C}_{k+1}}{\sqrt{k+1}} \leq \frac{\gamma^k\tilde{C}}{\sqrt{(k+1)!}} + \frac{\tilde{C}_k}{\sqrt{k}}\frac{4C}{\sqrt{k+1}}.$$

Let us define

$$A_k = \frac{\tilde{C}_k}{\sqrt{k}}, \quad M_{k+1} = \frac{\gamma^k\tilde{C}}{\sqrt{(k+1)!}}, \quad D_{k+1} = \frac{4C}{\sqrt{k+1}},$$

hence, we have that

$$A_{k+1} \leq M_{k+1} + A_kD_{k+1}.$$

By taking $\gamma$ such that $\gamma \geq 4C$ we obtain :

$$M_kD_{k+1} = \frac{4C}{\sqrt{k+1}}\frac{\tilde{C}\gamma^{k-1}}{\sqrt{k!}} \leq \frac{\gamma^k\tilde{C}}{\sqrt{(k+1)!}} = M_{k+1}.$$

Therefore, by induction

$$A_{k+1} \leq kM_{k+1} + \prod_{i=1}^{k+1}\frac{4C}{\sqrt{i}},$$

which implies that

$$\frac{\tilde{C}_{k+1}}{\sqrt{k+1}} \leq k\frac{\gamma^k}{\sqrt{(k+1)!}} + \frac{C^{k+1}4^{k+1}}{\sqrt{(k+1)!}}.$$

Then,

$$\tilde{C}_{k+1} \leq k\frac{\gamma^k}{\sqrt{k!}} + \frac{C^{k+1}4^{k+1}}{\sqrt{k!}}.$$

□

### *B.2.3* Proof of Lemma B.2.3

From the representation (B.2.24) of $[\partial/\partial x]p$ and from the arguments of the proof of Claims B.2.4, B.2.5 and B.2.6, we can see that for all $t \in [0,T]$, all $x, z \in \mathbb{R}^d$, the mapping $y \mapsto [\partial/\partial x]p(\bar{\mu}(t,x); t, y; T, z)$ is continuously differentiable and satisfies B.2.3. Also, from classical results on parametrix, and arguments of Claims B.2.4, B.2.5 and B.2.6 , we can see that for all $t \in [0,T]$, all $y, z \in \mathbb{R}^d$, the mapping $x \mapsto [\partial/\partial y]p(\bar{\mu}(t,x); t, y; T, z)$ is continuously differentiable and satisfies the bound of Lemma B.2.3. This concludes the proof.

## *B.3* Proof of Proposition B.1.2

First of all, we have from Sznitman (see [Szn91]) that

**Claim B.3.1.** *The system :*

$$\begin{cases} dX_t = \sum_{i=0}^{d} V_i(t, X_t, \mathbb{P}_{X_t}) dB_t^i \\ X_0 = x \end{cases} \tag{B.3.1}$$

*where the $V_i$, $i = 0, \cdots, d$ are supposed to be Lipschitz continuous in space and Lipschitz continuous in the McKean-Vlasov component w.r.t. the Wassertein-2 distance*[39] *admits a unique strong solution. Moreover, this solution is in $\mathbb{L}_p$ for any $p \geq 1$.*

And we prove that :

**Claim B.3.2.** *Let $(X_t^x)_{0 \leq t \leq T}$ be the solution of (B.1.1) starting from $x$. Let $\alpha_i^1, \alpha_i^2$ and $\beta_i$ be $3 \times d$ positive reals less than 1 such that the $V_i's$ are $C_b^{1+\alpha_i^1, 1+\alpha_i^2}$ and the $\varphi_i's$ are $C_b^{1+\beta_i}$, $i = 0, \cdots, d$.*

*Then, for all $t$ in $[0,T]$, the mapping $X_t :\ x \in \mathbb{R}^d \mapsto X_t^x$ is $\mathbb{L}_p$, $p \geq 2$ differentiable. By denoting by $Y_t^k$ its derivative in the direction $e_k$ (where $e_k$ is the $k^{th}$ vector of the canonical basis of $\mathbb{R}^d$) we have :*

$$\begin{cases} dX_t^x = \sum_{i=0}^{d} V_i\left(t, X_t^x, \mathbb{E}\left[\varphi_i(X_t^x)\right]\right) dB_t^i \\ dY_t^k = \sum_{i=0}^{d} \left\{ \partial_x V_i\left(t, X_t^x, \mathbb{E}\left[\varphi_i(X_t^x)\right]\right) Y_t^k + \partial_y V_i\left(t, X_t^x, \mathbb{E}\left[\varphi_i(X_t^x)\right]\right) \mathbb{E}\left[\partial_x \varphi_i(X_t^x) \cdot Y_t^k\right] \right\} dB_t^i \\ X_0^x = x,\ Y_0^k = 1 \end{cases} \tag{B.3.2}$$

*for all $k \in \{1, \cdots, d\}$ and where $\partial_x V_i$ is the matrix $\left([\partial V_i^l / \partial x_k]\right)_{1 \leq k,l \leq d}$.*

Then, the proof of Proposition B.1.2 follow by iterating these two claims.

### *B.3.1* Proof of Claim B.3.2

Thanks to Claim B.3.1, we know that the system (B.3.2) admits a unique solution which is $\mathbb{L}_2$ bounded. For any $k \in \{1, \cdots, d\}$, we denote by $e_k$ the $k^{\text{th}}$ vector of the canonical basis of $\mathbb{R}^d$. Let $\epsilon$ be a real positive number, we define :

$$Z_t^k = \frac{X_t^{x+\epsilon e_k} - X_t^x}{\epsilon} - Y_t^k. \tag{B.3.3}$$

39. For $\nu, \nu'$ two probability measures on $\mathbb{R}^d$ admitting moments of second order, we define by $d_2(\nu, \nu')$ the Wasserstein-2 distance given by $d_2(\nu, \nu') = \sqrt{\inf\{\int |x-y|^2 d\pi(x,y)\}}$, where the infimum is taken over all the probability measure $\pi$ admitting $\nu$ and $\nu'$ as marginals.

where the process $(Y_t^k)_{t\geq 0}$ is defined by (B.3.2).

$$
\begin{aligned}
dZ_t^k &= \sum_{i=0}^{d}\Big\{\frac{1}{\epsilon}\left(V_i\left(t, X_t^x, \mathbb{E}\left[\varphi_i(X_t^x)\right]\right) - V_i\left(t, X_t^{x+\epsilon e_k}, \mathbb{E}\left[\varphi_i(X_t^{x+\epsilon e_k})\right]\right)\right) \\
&\quad -\frac{\partial}{\partial x}V_i\left(t, X_t^x, \mathbb{E}\left[\varphi_i(X_t^x)\right]\right)Y_t^k - \frac{\partial}{\partial y}V_i\left(t, X_t^x, \mathbb{E}\left[\varphi_i(X_t^x)\right]\right)\mathbb{E}\left[\frac{\partial}{\partial x}\varphi_i(X_t^x)\cdot Y_t^k\right]\Big\}dB_t^i.
\end{aligned}
$$

Note that Claim B.3.1 ensures the existence of a unique $Y^k$ and $Z$. From Mean Value Theorem (MVT), the process $Z$ can be written as the solution of :

$$
\begin{aligned}
Z_t^k &= \sum_{i=0}^{d}\Big\{\int_0^t\int_0^1 \frac{\partial}{\partial x}V_i\left(s, X_s^x + v\epsilon(Z_s^k+Y_s^k), \mathbb{E}\left[\varphi_i(X_s^x + v\epsilon(Z_s^k+Y_s^k))\right]\right)dv(Z_s^k+Y_s^k)dB_s^i \\
&\quad + \int_0^t\int_0^1 \frac{\partial}{\partial y}V_i\left(s, X_s^x v\epsilon(Z_s^k+Y_s^k), \mathbb{E}\left[\varphi_i(X_s^x + v\epsilon(Z_s^k+Y_s^k))\right]\right) \\
&\quad \times\mathbb{E}\left[\frac{\partial}{\partial x}\varphi_i((X_s^x + v\epsilon(Z_s^k+Y_s^k))\cdot(Z_s^k+Y_s^k)\right]dvdB_s^i\Big\} \\
&\quad - \sum_{i=0}^{d}\int_0^t \frac{\partial}{\partial x}V_i\left(s, X_s^x, \mathbb{E}\left[\varphi_i(X_s^x)\right]\right)Y_s^k + \frac{\partial}{\partial y}V_i\left(s, X_s^x, \mathbb{E}\left[\varphi_i(X_s^x)\right]\right)\mathbb{E}\left[\frac{\partial}{\partial x}\varphi_i(X_s^x)\cdot Y_s^k\right]dB_s^i.
\end{aligned}
$$

So that

$$
\begin{aligned}
& Z_t^k \\
&= \sum_{i=0}^{d}\Big\{\int_0^t\int_0^1\Big[\frac{\partial}{\partial x}V_i\left(s, X_s^x + v\epsilon(Z_s^k+Y_s^k), \mathbb{E}\left[\varphi_i(X_s^x + v\epsilon(Z_s^k+Y_s^k))\right]\right) && \text{(B.3.4)}\\
&\quad -\frac{\partial}{\partial x}V_i\left(s, X_s^x, \mathbb{E}\left[\varphi_i(X_s^x)\right]\right)\Big]dvY_s^kdB_s^i \\
&+ \int_0^t\int_0^1\Big[\frac{\partial}{\partial y}V_i\left(s, X_s^x + v\epsilon(Z_s^k+Y_s^k), \mathbb{E}\left[\varphi_i(X_s^x + v\epsilon(Z_s^k+Y_s^k))\right]\right) && \text{(B.3.5)}\\
&\qquad \times\mathbb{E}\left[\frac{\partial}{\partial x}\varphi_i((X_s^x + v\epsilon(Z_s^k+Y_s^k))\cdot Y_s^k\right] \\
&\quad -\frac{\partial}{\partial y}V_i\left(s, X_s^x, \mathbb{E}\left[\varphi_i(X_s^x)\right]\right)\mathbb{E}\left[\frac{\partial}{\partial x}\varphi_i(X_s^x)\cdot Y_s^k\right]\Big]dvdB_s^i \\
&+ \int_0^t\int_0^1 \frac{\partial}{\partial x}V_i\left(s, X_s^x + v\epsilon(Z_s^k+Y_s^k), \mathbb{E}\left[\varphi_i(X_s^x + v\epsilon(Z_s^k+Y_s^k))\right]\right)dv && \text{(B.3.6)}\\
&\quad \times Z_s^kdB_s^i && \text{(B.3.7)}\\
&+ \int_0^t\int_0^1 \frac{\partial}{\partial y}V_i\left(s, X_s^x + v\epsilon(Z_s^k+Y_s^k), \mathbb{E}\left[\varphi_i(X_s^x + v\epsilon(Z_s^k+Y_s^k))\right]\right) && \text{(B.3.8)}\\
&\qquad \times\mathbb{E}\left[\frac{\partial}{\partial x}\varphi_i((X_s^x + v\epsilon(Z_s^k+Y_s^k))\cdot Z_s^k\right]dvdB_s^i\Big\}.
\end{aligned}
$$

For all $i \in \{1, \cdots, d\}$, by taking the square, the supremum, and the expectation of terms (B.3.7) and (B.3.8) one gets, from Cauchy-Schwarz and BDG inequality :

$$\mathbb{E}\left[\sup_{t\leq T} |(B.3.7) + (B.3.8)|^2\right] \leq C_T \int_0^T \mathbb{E} \sup_{u\leq s} |Z_u^k|^2 ds, \tag{B.3.9}$$

where $C_T$ depends on $T$.

Now, we focus on (B.3.5). For each $i \in \{1, \cdots, d\}$, it can be written as :

$$\begin{aligned}
&\int_0^t \int_0^1 \left[\frac{\partial}{\partial y} V_i\left(s, X_s^x + v\epsilon(Z_s^k + Y_s^k), \mathbb{E}\left[\varphi_i(X_s^x + v\epsilon(Z_s^k + Y_s^k))\right]\right)\right. \\
&\qquad \times \mathbb{E}\left[\frac{\partial}{\partial x}\varphi_i((X_s^x + v\epsilon(Z_s^k + Y_s^k)) \cdot Y_s^k\right] \\
&\quad \left. -\frac{\partial}{\partial y} V_i\left(s, X_s^x, \mathbb{E}\left[\varphi_i(X_s^x)\right]\right) \mathbb{E}\left[\frac{\partial}{\partial x}\varphi_i(X_s^x) \cdot Y_s^k\right]\right] dv dB_s^i \\
&= \int_0^t \int_0^1 \left\{\left[\frac{\partial}{\partial y} V_i\left(s, X_s^x + v\epsilon(Z_s^k + Y_s^k), \mathbb{E}\left[\varphi_i(X_s^x + v\epsilon(Z_s^k + Y_s^k))\right]\right)\right.\right. \\
&\qquad \left.\left. -\frac{\partial}{\partial y} V_i\left(s, X_s^x, \mathbb{E}\left[\varphi_i(X_s^x)\right]\right)\right] \mathbb{E}\left[\frac{\partial}{\partial x}\varphi_i((X_s^x + v\epsilon(Z_s^k + Y_s^k)) \cdot Y_s^k\right]\right\} dv dB_s^i \\
&\quad + \int_0^t \int_0^1 \frac{\partial}{\partial y} V_i\left(s, X_s^x, \mathbb{E}\left[\varphi_i(X_s^x)\right]\right) \\
&\qquad \times \mathbb{E}\left[\left[\frac{\partial}{\partial x}\varphi_i(X_s^x + v\epsilon(Z_s^k + Y_s^k)) - \frac{\partial}{\partial x}\varphi_i(X_s^x)\right] \cdot Y_s^k\right] dv dB_s^i
\end{aligned}$$

By taking the square, the supremum, and the expectation, we have from BDG and Cauchy-Schwarz inequality :

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\leq T} |(B.3.5)|^2\right] \quad \leq \quad & C\left(\int_0^T \sum_{j=1}^2 \epsilon^{2\alpha_i^j} \mathbb{E} \sup_{u\leq s} |Z_u^k + Y_u^k|^{2\alpha_i^j} + \epsilon^{2\beta_i} \mathbb{E} \sup_{u\leq s} |Z_u^k + Y_u^k|^{2\beta_i} ds\right) \\
& \times \int_0^T \mathbb{E} \sup_{u\leq s} |Y_u^k|^2 ds,
\end{aligned} \tag{B.3.10}$$

where $C$ depends on $T$.

Same arguments lead to :

$$\mathbb{E}\left[\sup_{t\leq T} |(B.3.4)|^2\right] \leq C \sum_{j=1}^2 \epsilon^{2\alpha_i^j} \left(\int_0^T \mathbb{E} \sup_{u\leq s} |Z_u^k + Y_u^k|^{2\alpha_i^j} ds\right) \int_0^T \mathbb{E} \sup_{u\leq s} |Y_u^k|^2 ds, \tag{B.3.11}$$

for $i \in \{1, \cdots, d\}$ and where $C$ depends on $T$.

Combining (B.3.9), (B.3.10) and (B.3.11) we obtain :

$$\mathbb{E}\left[\sup_{t\leq T} \left|Z_t^k\right|^2\right] \leq C\left\{\rho(\epsilon, Y, X, T) \int_0^T \mathbb{E} \sup_{u\leq s} |Y_u^k|^2 ds + \int_0^T \mathbb{E} \sup_{u\leq s} |Z_u^k|^2 ds\right\}, \tag{B.3.12}$$

where $\rho(\epsilon, Y, X, T)$ tends to 0 when $\epsilon$ tends to 0. Indeed, we have that for all $t$ in $[0, T]$

$$\epsilon^2 \mathbb{E}|Z_t^k + Y_t^k|^2 = \mathbb{E}|X_t^{x+\epsilon e_k} - X_t^x|^2.$$

By taking the square, the supremum, and the expectation of (B.3.4) and (B.3.5), we have from BDG and Cauchy-Schwarz inequality and estimate (B.3.9) that

$$\mathbb{E}[\sup_{t \le T} |Z_t^k|^2] \le C,$$

so that

$$\mathbb{E}\left[\sup_{t \le T} |X_t^{x+\epsilon e_k} - X_t^x|^2\right] \to 0,$$

when $\epsilon \to 0$.

Now we come back to (B.3.12) : by Gronwall's Lemma and by letting $\epsilon$ tends to 0, we can conclude that for all $k$ in $\{1, \cdots, d\}$, the $\mathbb{L}_2$ limit of the supremum of $Z_t^k$ over $[0, T]$ is equal to 0. This proves Claim B.3.2. □

# Bibliographie


[AKH02] F. Antonelli and A. Kohatsu-Higa. Rate of convergence of a particle method to the solution of the McKean–Vlasov equation. *The Annals of Applied Probability*, 12(2) :423–476, May 2002. Mathematical Reviews number (MathSciNet) : MR1910635 ; Zentralblatt MATH identifier : 1015.60048.

[Bas98] R. F. Bass. *Diffusions and elliptic operators.* Probability and its Applications (New York). Springer-Verlag, New York, 1998.

[BDH+03] P. Briand, B. Delyon, Y. Hu, E. Pardoux, and L. Stoica. Lp solutions of backward stochastic differential equations. *Stochastic Processes and their Applications*, 108(1) :109–129, November 2003.

[Ber09] L. Bergomi. Smile dynamics III. SSRN Scholarly Paper ID 1493308, Social Science Research Network, Rochester, NY, October 2009.

[BHH+11] N. Bush, B. M. Hambly, H. Haworth, L. Jin, and C. Reisinger. Stochastic evolution equations in portfolio credit modelling. *SIAM Journal on Financial Mathematics*, 2(1) :627–664, January 2011.

[BJPR09] T. R. Bielecki, M. Jeanblanc-Picqué, and M. Rutkowski. *Credit Risk Modeling.* Osaka University Press Osaka, 2009.

[BLP09] R. Buckdahn, J. Li, and S. Peng. Mean-field backward stochastic differential equations and related partial differential equations. *Stochastic Processes and their Applications*, 119(10) :3133–3154, 2009.

[Bos05] M. Bossy. Some stochastic particle methods for nonlinear parabolic PDEs. In *GRIP—Research Group on Particle Interactions*, volume 15 of *ESAIM Proc.*, page 18–57. EDP Sci., Les Ulis, 2005.

[BT96] V. Bally and D. Talay. The law of the euler scheme for stochastic differential equations. i. convergence rate of the distribution function. *Probability Theory and Related Fields*, 104(1) :43–60, 1996.

[BT08] C. Bayer and J. Teichmann. Cubature on wiener space in infinite dimension. *Proceedings of the Royal Society A : Mathematical, Physical and Engineering Science*, 464(2097) :2493–2516, September 2008.

[BZ11] M. Bramanti and M. Zhu. L^p and schauder estimates for nonvariational operators structured on hörmander vector fields with drift. arXiv e-print 1103.5116, March 2011.

[Car10] P. Cardaliaguet. Notes on mean field games. Technical report, 2010.

[CD11] D. Crisan and F. Delarue. Sharp gradient bounds for solutions of degenerate semi-linear partial differential equations. June 2011.

[CD12a] R. Carmona and F. Delarue. Mean field forward-backward stochastic differential equations. Technical report, 2012.

[CD12b] R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. Technical report, 2012.

[CD12c] D. Crisan and F. Delarue. Sharp derivative bounds for solutions of degenerate semi-linear partial differential equations. *Journal of Functional Analysis*, 263(10) :3024–3101, 2012.

[CDL12] R. Carmona, F. Delarue, and A. Lachapelle. Control of McKean–Vlasov dynamics versus mean field games. *Mathematics and Financial Economics*, pages

1–36, 2012.

[CdR12] P. E. Chaudru de Raynal. Regularization by stochastic drift. arXiv e-print 1205.6688, May 2012.

[CdRGT13] P. E. Chaudru de Raynal and C. A. Garcia Trillos. A cubature based algorithm to solve decoupled McKean-Vlasov forward backward stochastic differential equations. arXiv e-print 1307.6427, July 2013.

[CG07] D. Crisan and S. Ghazali. On the convergence rates of a general class of weak approximations of SDEs. In *Stochastic differential equations : theory and applications*, volume 2 of *Interdiscip. Math. Sci.*, page 221–248. World Sci. Publ., Hackensack, NJ, 2007.

[CM02] P. Cattiaux and L. Mesnager. Hypoelliptic non-homogeneous diffusions. *Probability Theory and Related Fields*, 123(4) :453–483, August 2002.

[CM10a] D. Crisan and K. Manolarakis. Second order discretization of backward SDEs. *arXiv :1012.5650*, December 2010.

[CM10b] D. Crisan and K. Manolarakis. Solving backward stochastic differential equations using the cubature method. application to nonlinear pricing. In *Progress in analysis and its applications*, page 389–397. World Sci. Publ., Hackensack, NJ, 2010.

[CM12] D. Crisan and K. Manolarakis. Solving backward stochastic differential equations using the cubature method : application to nonlinear pricing. *SIAM Journal on Financial Mathematics*, 3 :534–571, 2012.

[Dav07] A. M. Davie. Uniqueness of solutions of stochastic differential equations. *International Mathematics Research Notices. IMRN*, (24) :Art. ID rnm124, 26, 2007.

[DFP06] M. Di Francesco and S. Polidoro. Schauder estimates, harnack inequality and gaussian lower bound for kolmogorov-type operators in non-divergence form. *Advances in Differential Equations*, 11(11) :1261–1320, 2006.

[DL89] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and sobolev spaces. *Inventiones Mathematicae*, 98(3) :511–547, 1989.

[DM06] F. Delarue and S. Menozzi. A forward-backward stochastic algorithm for quasi-linear PDEs. *The Annals of Applied Probability*, 16(1) :140–184, 2006.

[DM10] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *Journal of Functional Analysis*, 259(6) :1577–1630, 2010.

[DTV12] P. Doersek, J. Teichmann, and D. Veluscek. Cubature methods for stochastic (partial) differential equations in weighted spaces. *arXiv :1201.4024*, January 2012.

[FF11] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics. An International Journal of Probability and Stochastic Processes*, 83(3) :241–257, 2011.

[FGP10] F. Flandoli, M. Gubinelli, and E. Priola. Well-posedness of the transport equation by stochastic perturbation. *Inventiones Mathematicae*, 180(1) :1–53, 2010.

[Fla11] F. Flandoli. *Random perturbation of PDEs and fluid dynamic models*, volume 2015 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010.

[Fri64] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.

[Fri06] A. Friedman. *Stochastic differential equations and applications.* Dover Publications Inc., Mineola, NY, 2006. Two volumes bound as one, Reprint of the 1975 and 1976 original published in two volumes.

[Fri08] A. Friedman. *Partial Differential Equations of Parabolic Type.* Dover Publications, April 2008.

[FS06] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer New York, 2006.

[GHL11] J. Guyon and P. Henry-Labordere. The smile calibration problem solved. SSRN Scholarly Paper ID 1885032, Social Science Research Network, Rochester, NY, July 2011.

[GL11] L. G. Gyurkó and T. Lyons. Efficient and practical implementations of cubature on wiener space. In *Stochastic analysis 2010*, page 73–111. Springer, Heidelberg, 2011.

[HDK+84] P. L. Hennequin, R. M. Dudley, H. Kunita, F. Ledrappier, and École d'été de probabilités. *Ecole d'été de probabilités de Saint-Flour XII-1982.* Springer-Verlag, Berlin, 1984.

[Hö67] L. Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, 119 :147–171, 1967.

[IDR10] P. Imkeller and G. Dos Reis. Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stochastic Processes and their Applications*, 120(3) :348–379, March 2010.

[Kol34] A. Kolmogorov. Zufällige bewegungen. (zur theorie der brownschen bewegung). *Ann. of Math., II. Ser.*, 35 :116–117, 1934.

[KP92] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York).* Springer-Verlag, Berlin, 1992.

[KR05] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probability Theory and Related Fields*, 131(2) :154–196, 2005.

[Kry69] N. V. Krylov. Itô's stochastic integral equations. *Akademija Nauk SSSR. Teorija Verojatnoste\uı\ i ee Primenenija*, 14 :340–348, 1969.

[Kry09] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A. B. Aries, Reprint of the 1980 edition.

[KS79] N. V. Krylov and M. V. Safonov. An estimate for the probability of a diffusion process hitting a set of positive measure. *Doklady Akademii Nauk SSSR*, 245(1) :18–20, 1979.

[KS84] S. Kusuoka and D. Stroock. Applications of the malliavin calculus. i. In *Stochastic analysis (Katata/Kyoto, 1982)*, volume 32 of *North-Holland Math. Library*, page 271–306. North-Holland, Amsterdam, 1984.

[KS85] S. Kusuoka and D. W. Stroock. Applications of the malliavin calculus. II. *Journal of the Faculty of Science. University of Tokyo. Section IA. Mathematics*, 32(1) :1–76, 1985.

[KS87] S. Kusuoka and D. Stroock. Applications of the malliavin calculus. III. *Journal of the Faculty of Science. University of Tokyo. Section IA. Mathematics*, 34(2) :391–442, 1987.

[Kun82] H. Kunita. *Stochastic differential equation and stochastic flows of diffeomorphisms.* Springer-Verlag, 1982.

[Kus01] S. Kusuoka. Approximation of expectation of diffusion process and mathematical finance. In *Taniguchi Conference on Mathematics Nara '98*, volume 31 of *Adv. Stud. Pure Math.*, page 147–165. Math. Soc. Japan, Tokyo, 2001.

[LL06a] J. M. Lasry and P.L. Lions. Jeux à champ moyen. i. le cas stationnaire. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 343(9) :619–625, 2006.

[LL06b] J. M. Lasry and P.L. Lions. Jeux à champ moyen. II. horizon fini et contrôle optimal. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 343(10) :679–684, 2006.

[LL07] J.M. Lasry and P.L. Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1) :229–260, 2007.

[LL12] C. Litterer and T. Lyons. High order recombination and an application to cubature on wiener space. *The Annals of Applied Probability*, 22(4) :1301–1327, August 2012. Zentralblatt MATH identifier : 06083936 ; Mathematical Reviews number (MathSciNet) : MR2985162.

[LV04] T. Lyons and N. Victoir. Cubature on wiener space. *Proceedings of the Royal Society of London. Series A : Mathematical, Physical and Engineering Sciences*, 460(2041) :169–198, January 2004.

[Ma02] J. Ma. Representation theorems for backward stochastic differential equations. *The Annals of Applied Probability*, 12(4) :1390–1418, November 2002. Mathematical Reviews number (MathSciNet) : MR1936598 ; Zentralblatt MATH identifier : 1017.60067.

[Men11] S. Menozzi. Parametrix techniques and martingale problems for some degenerate kolmogorov equations. *Electronic Communications in Probability*, 16 :234–250, 2011.

[MS67] H. P. McKean, Jr. and I. M. Singer. Curvature and the eigenvalues of the laplacian. *Journal of Differential Geometry*, 1(1) :43–69, 1967.

[Nee11] C. Nee. *Sharp Gradient Bounds for the Diffusion Semigroup.* PhD thesis, Imperial College London, 2011.

[Pha09] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer, 2009.

[PP90] É. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1) :55–61, 1990.

[PP92] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lecture Notes in Control and Inform. Sci.*, page 200–217. Springer, Berlin, 1992.

[Pro04] P.E. Protter. *Stochastic Integration and Differential Equation.* Springer, 2004.

[Ric11] A. Richou. Numerical simulation of BSDEs with drivers of quadratic growth. *The Annals of Applied Probability*, 21(5) :1933–1964, 2011.

[Str08] D. W. Stroock. *Partial differential equations for probabilists*, volume 112 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2008.

[SV79] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979.

[Szn91] A. S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, page

165–251. Springer, Berlin, 1991.

[Tei06] J. Teichmann. Calculating the greeks by cubature formulae. *Proceedings of The Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences*, 462(2066) :647–670, 2006.

[TV03] D. Talay and O. Vaillant. A stochastic particle method with random weights for the computation of statistical solutions of McKean-Vlasov equations. *The Annals of Applied Probability*, 13(1) :140–180, January 2003. Mathematical Reviews number (MathSciNet) : MR1951996 ; Zentralblatt MATH identifier : 1026.60110.

[Ver80] A. Ju. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Matematicheski\uı\ Sbornik. Novaya Seriya*, 111(153)(3) :434–452, 480, 1980.

[Ver83] A. Yu. Veretennikov. Stochastic equations with diffusion that degenerates with respect to part of the variables. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 47(1) :189–196, 1983.

[YW71] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University*, 11 :155–167, 1971.

[YZ99] J. Young and X. Y. Zhou. *Stochastic controls : hamiltonian systems and HJB equations*. Springer, New York [etc.], 1999.

[Zha04] J. Zhang. A numerical scheme for BSDEs. *The Annals of Applied Probability*, 14(1) :459–488, 2004.

[Zha05] X. Zhang. Strong solutions of SDES with singular drift and sobolev diffusion coefficients. *Stochastic Processes and their Applications*, 115(11) :1805–1818, 2005.

[Zvo74] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135) :129–149, 152, 1974.

**Équations différentielles stochastiques : résolubilité forte d'équations singulières dégénérées ; analyse numérique de systèmes progressifs-rétrogrades de McKean-Vlasov**

**Résumé.** Cette thèse traite de deux sujets : la résolubilité forte d'équations différentielles stochastiques à dérive hölderienne et bruit hypoelliptique et la simulation de processus progressifs-rétrogrades découplés de McKean-Vlasov.

Dans le premier cas, on montre qu'un système hypoelliptique, composé d'une composante diffusive et d'une composante totalement dégénérée, est fortement résoluble lorsque l'exposant de la régularité Hölder de la dérive par rapport à la composante dégénérée est strictement supérieur à $2/3$. Ce travail étend au cadre dégénéré les travaux antérieurs de Zvonkin (1974), Veretennikov (1980) et Krylov et Röckner (2005). L'apparition d'un seuil critique pour l'exposant peut-être vue comme le prix à payer pour la dégénérescence. La preuve repose sur des résultats de régularité de la solution de l'EDP associée, qui est dégénérée, et est basée sur une méthode parametrix.

Dans le second cas, on propose un algorithme basé sur les méthodes de cubature pour la simulation de processus progessifs-rétrogrades découplés de McKean-Vlasov apparaissant dans des problèmes de contrôle dans un environnement de type champ moyen. Cet algorithme se divise en deux parties. Une première étape de construction d'un arbre de particules, à dynamique déterministe, approchant la loi de la composante progressive. Cet arbre peut être paramétré de manière à obtenir n'importe quel ordre d'approximation (en terme de pas de discrétisation de l'intervalle). Une seconde étape, conditionnelle à l'arbre, permettant l'approximation de la composante rétrograde. Deux schémas explicites sont proposés permettant un ordre d'approximation de 1 et 2.

**Mots Clés : analyse stochastique, EDS, EDP, résolubilité forte, parametrix, dégénérescence, EDSPR, schéma numérique, algorithme, Mckean-Vlasov, Cubature, champ moyen**

---

**Stochastic differential equations : strong well-posedness of singular and degenerate equations ; numerical analysis of decoupled forward backward systems of McKean-Vlasov type.**

**Abstract.** This thesis deals with two subjects : the strong well-posedness of stochastic differential equations with Hölder drift and hypoelliptic noise and the simulation of decoupled forward backward stochastic differential equations of McKean-Vlasov type.

In the first work, we study a class of degenerate system with hypoelliptic noise. We prove that strong well-posedness holds for this system when the drift is only Hölder, with Hölder exponent larger than the critical value $2/3$. This work extends to the degenerate setting the earlier results obtained by Zvonkin (1974), Veretennikov (1980) and Krylov and Röckner (2005). The existence of a threshold for the Hölder exponent in the degenerate case may be understood as the price to pay to balance the degeneracy of the noise. Our proof relies on regularization properties of the associated PDE, which is degenerate in the current framework and is based on a parametrix method.

In the second work, we propose a new algorithm to approach weakly the solution of a McKean-Vlasov stochastic differential equation. Based on the cubature method, the algorithm is deterministic differing from the usual methods based on interacting particles. It can be parametrized in order to obtain a given order of convergence.

Then, we construct implementable algorithms to solve decoupled forward backward stochastic differential equations of McKean-Vlasov type, which appear in some stochastic control problems in a mean field environment. We give two algorithms and show that they have convergence of orders one and two under appropriate regularity conditions.

**Keywords : stochastic analysis, SDE, PDE, strong well-posedness, degeneracy, parametrix, FBSDE, numerical scheme, algorithm, McKean-Vlasov, Cubature, Mean Field**